

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 17, 1999, 03:08:16 ; Search time 64.1 Seconds
(without alignments)
11.086 Million cell updates/sec

Title: US-09-037-657-32

Perfect score: 139

Sequence: 1 MYLASSSTSIHMTLLMLLMLFHLGLQASIS 30

Scoring table: BLOSUM62

Searched: 188963 seqs, 2366106 residues

Database: A.Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	139	100.0	166	1	P50161	Sequence encoded b
2	139	100.0	30	1	M10638	Murine IL-3 signal
3	139	100.0	30	1	M09823	Interleukin-3 sign
4	139	100.0	359	1	M56260	Construct contai
5	139	100.0	30	1	M55017	Murine IL3 signal
6	48	34.5	694	1	M31267	Drosophila fritzle
7	46	33.1	116	1	R71911	Anther specific pro
8	45	33.4	96	1	M46836	Bacillus thuringie
9	45	33.4	652	1	M30637	Human 7-transmembr
10	45	32.4	521	1	M30638	Partial human 7-tr
11	43	30.9	1268	1	M58774	Human breast cance
12	43	30.9	481	1	M55029	G-protein coupled
13	43	30.9	542	1	M55030	G-protein coupled
14	42	30.2	266	1	M71460	Protein encoded by
15	42	30.2	306	1	M90236	E. coli antibiotic
16	42	30.2	299	1	M90237	E. coli antibiotic
17	42	30.2	266	1	M90238	E. coli antibiotic
18	42	30.2	51	1	Y02760	Human secreted pro
19	42	30.2	55	1	Y12970	Protein sequence o
20	41.5	29.9	428	1	M63713	Human hiki protein
21	41.5	29.9	427	1	M63717	Human hiki protein
22	41	29.5	315	1	R43568	Mouse MSH-R. Melan
23	41	29.5	709	1	R99801	Mouse MSH-R. Melan
24	41	29.5	315	1	M37826	Mouse melanocyte s
25	41	29.5	315	1	M87864	Mouse melanocyte s
26	40	28.8	336	1	R28149	Sugar beet beta-1,
27	40	28.8	1276	1	R35199	Mouse multiling re
28	40	28.8	30	1	R63638	Sheep erythropoiet
29	40	28.8	22	1	R65246	N-terminal of 48 k
30	40	28.8	308	1	R75757	BAV3 ORF2 product.
31	40	28.8	523	1	M78915	Mouse major periph
32	40	28.8	248	1	M94567	Mouse major periph
33	40	28.8	45	1	Y12940	Amnio acid sequenc
34	39.5	28.4	132	1	P70411	ORF 8 gene product
35	39.5	28.4	244	1	M29149	Human high affinity
36	39.5	28.4	244	1	M75918	Human beta subunit
37	39	28.1	205	1	R05232	Human lymphotoxin
38	39	28.1	1093	1	R56979	Human myotonic dys
39	39	28.1	1206	1	R82672	CD5 associating p
40	39	28.1	1989	1	R99640	Peripheral nervous
41	39	28.1	1977	1	R99641	Peripheral nervous
42	39	28.1	467	1	M00382	Bacillus cellulase
43	39	28.1	467	1	M05731	Cellulase. Cellula

ALIGNMENTS

44	39	28.1	81	1	M30082	Rat perlephn. GDN
45	39	28.1	164	1	M38705	S. pneumoniae phos

RESULT 1	
ID	P50161 standard; Protein; 166 AA.
AC	P50161;
DT	27-NOV-1991 (first entry)
DE	Sequence encoded by cDNA clone exhibiting multi-lineage cellular
DE	growth factor activity.
KW	Growth factor; multi-lineage; mast cell; haematopoietic.
OS	Mus musculus.
PN	Ep-138133-A.
PD	24-APR-1985.
PF	29-SEP-1984; 11677.
PR	04-OCT-1983; US-539050.
PR	19-MAR-1984; US-530867.
PA	(SCHE) SCHERING CORP.
PA	(DNAX-) DNAX RES INST MOLEC.
PI	Yokota T, Lee PD, Rennick DW, Aral KI.
DR	WPI: 85-100349/17.
DR	N-PSDB; N50199.
PT	New poly:peptide(s) having growth factor activities - are prepd.
PT	by recombinant DNA procedures
PS	Claim 5; Fig 1; 64pp; English.
CC	The cDNA is derived from messenger RNA isolated from a mouse T-cell
CC	line after activation with concanavalin A. The cDNA was cloned by
CC	incorporation into a plasmid vector, which then transformed into
CC	E.coli. The plasmid vector also contained DNA segments from the SV40
CC	virus, permitting expression of the cDNA after transfection into a
CC	mammalian host cell, such as monkey COS-7 cells. The polypeptide
CC	includes a potential leader sequence of about 19 AAs.
SO	Sequence 166 AA.

Query Match	100.0%; Score 139; DB 1; Length 166;
Best Local Similarity	100.0%; Pred. No. 7.2e-14;
Matches	30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1 MYLASSSTSIHMTLLMLLMLFHLGLQASIS 30
DB	1 MYLASSSTSIHMTLLMLLMLFHLGLQASIS 30

RESULT 2	
ID	M10638 standard; Peptide; 30 AA.
AC	M10638;
DT	23-JUN-1997 (first entry)
DE	Murine IL-3 signal sequence.
KW	LEK; ligand for eph-related kinase; ERK; NLERK2;
KW	receptor protein tyrosine kinase; cell proliferation;
KW	cell differentiation; cell survival; nerve cell; interleukin-3;
OS	Mus sp.
PN	M09704091-AL.
PD	06-FEB-1997.
PF	19-JUL-1996; AD0460.
PR	20-JUL-1995; AU-004263.
PR	27-NOV-1995; AU-006847.
PR	22-DEC-1995; AU-007299.
PR	05-FEB-1996; AU-007890.
PA	(AMRA-) AMRAD OPERATIONS PTY LTD.
PI	Nicola NA;
DR	WPI: 97-132632/12.
PT	Nucleic acid mol. encoding ligand for eph-related kinase - useful
PT	for treatment of, pref. neuronal, cells to increase survival,
PT	proliferation and differentiation.
PS	Example 3; Page 30; 71pp; English.

Sequence Comparison

CC A peptide sequence (W10638) comprises the signal sequence of
 CC mouse interleukin-3 (IL-3). Mammalian expression vector pEF-BOS
 CC was engineered to contain DNA for the IL-3 signal sequence and for
 CC a FLAG epitope. PCR fragments (see also T60970-72) coding for
 CC mature or soluble NLERK2 (see also W10637), a novel ligand for
 CC eph kinase (LEK), were cloned into the vector to allow prodn. of
 CC recombinant NLERK2 proteins in transfected COS cells.
 SQ Sequence 30 AA;

Query Match 100.0%; Score 139; DB 1; Length 30;
 Best Local Similarity 100.0%; Pred. No. 1e-14;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MYLASSSTSIHTMLLMLFHLGLQASIS 30
 DB 1 MYLASSSTSIHTMLLMLFHLGLQASIS 30

RESULT 3

ID W09823 standard; Peptide; 30 AA.
 AC W09823;
 DT 15-JUL-1997 (first entry)
 DE Interleukin-3 signal peptide.
 KW Interleukin-3; IL-3; signal peptide; NR4; haemopoietin receptor;
 KW Interleukin-13 receptor; cytokine; allergy; asthma; therapy.
 OS Mus sp.
 PN W09715663-A1.
 PD 01-MAY-1997.
 PF 23-OCT-1996; AU0668.
 PR 23-OCT-1995; AU-006135.
 PR 22-DEC-1995; AU-007276.
 PR 09-SEP-1996; AU-002208.
 PA (AMRA-) AMRAD OPERATIONS PTY LTD.
 PI Hilton DJ, Metcalf D, Nicola NA, Willson T, Zhang JG;
 DR WPI; 97-259018/23.
 PT DNA encoding animal haemopoietin receptor which interacts with
 PT interleukin-13 - useful to treat asthma, allergy or condition
 PT exacerbated by IgE production.
 PS Example 2; Page 56; 93pp; English.
 CC Using PCR, a derivative of novel haemopoietin receptor NR4 cDNA was
 CC generated which encoded the interleukin-3 signal peptide (W09823)
 CC and an N-terminal FLAG epitope tag (W09824) preceding the mature
 CC coding region (Thr27-Pro242) of murine NR4 (see also W09821). The
 CC PCR product was cloned into the mammalian expression vector
 CC pEF-BOS.
 SQ Sequence 30 AA;

Query Match 100.0%; Score 139; DB 1; Length 30;
 Best Local Similarity 100.0%; Pred. No. 1e-14;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MYLASSSTSIHTMLLMLFHLGLQASIS 30
 DB 1 MYLASSSTSIHTMLLMLFHLGLQASIS 30

RESULT 4
 W56260 standard; Protein; 359 AA.

AC W56260;
 DT 16-SEP-1998 (first entry)
 DE Construct containing mature interleukin-13 binding protein.
 KW Therapeutic; IL-3 mediated condition; allergy; asthma; diagnosis;
 KW autoimmune disease; antibody; immunotherapy.
 OS Homo sapiens.
 PN W09810638-A1.
 PD 19-MAR-1998.
 PF 10-SEP-1997; AU0591.
 PR 27-FEB-1997; AU-005374.
 PR 10-SEP-1996; AU-002262.

PA (AMRA-) AMRAD OPERATIONS PTY LTD.
 PI Hilton DJ, Nicola NA, Simpson RJ, Zhang J;
 DR WPI; 98-207062/18.
 DR N-PSDB; V22701.

PT New isolated interleukin-13 binding protein - used to develop
 PT products for therapy e.g. for allergic conditions such as asthma or
 PT for diagnosis or detection.
 PS Example 14; Page 52-53; 69pp; English.
 CC The IL-13 binding protein and related therapeutic molecules can be used
 CC in the antagonism of at least one IL-13 activity. They can be used for
 CC treating IL-13 mediated conditions such as certain allergic conditions
 CC such as asthma or to inactivate locally administered IL-13 after IL-13
 CC treatment. The products can also be used as diagnostic agents, e.g. for
 CC detecting autoimmune diseases. The antibodies can also be used for
 CC immunotherapy and may also be used as a diagnostic tool.
 SQ Sequence 359 AA;

Query Match 100.0%; Score 139; DB 1; Length 359;
 Best Local Similarity 100.0%; Pred. No. 1.7e-13;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MYLASSSTSIHTMLLMLFHLGLQASIS 30
 DB 1 MYLASSSTSIHTMLLMLFHLGLQASIS 30

RESULT 5

ID W55017 standard; Protein; 30 AA.
 AC W55017;
 DT 29-SEP-1998 (first entry)
 DE Murine IL3 signal sequence.
 KW Haemopoietin receptor; cell proliferation; cell differentiation; cancer;
 KW cell survival; therapeutic; neuronal proliferation; drug screening;
 KW Mouse.
 OS Mus sp.
 PN W09811225-A2.
 PD 19-MAR-1998.
 PF 11-SEP-1997; G02479.
 PR 11-SEP-1996; AU-002246.
 PA (AMRA-) AMRAD OPERATIONS PTY LTD.
 PI (DZIE-) DZIELEMSKA H E.
 PI Alexander W, Fabril L, Farley A, Hilton DJ, Kikuchi Y,
 PI Kojima T, Maeda M, Nash A, Nicola NA, Raker S, Willson T,
 PI Zhang J;
 DR WPI; 98-260970/23.
 PT New isolated haemopoietin receptor - used for developing products
 PT for modulating proliferation, differentiation and survival of cells,
 PT e.g. neuronal cells.
 PS Claim 29(1); Page 54; 182pp; English.
 CC The mouse IL3 signal sequence was used to study the expression of mouse
 CC NR6.1. NR6.1 is a form of NR6 a novel Haemopoietin receptor (HR).
 CC Interaction between the novel HR and a ligand facilitates proliferation,
 CC differentiation and survival of a wide variety of cells. The HR and it's
 CC derivatives can be used for modulating the activity of the receptors e.g.
 CC to regulate development, maintenance or regeneration in an array of
 CC different cells and tissues in vitro and in vivo. They can be present in
 CC therapeutics used for modulating neuronal proliferation, differentiation
 CC and survival. The products can also be used for detection and diagnosis,
 CC e.g. for cancers or predisposition to cancers, or for drug screening.
 SQ Sequence 30 AA;

Query Match 100.0%; Score 139; DB 1; Length 30;
 Best Local Similarity 100.0%; Pred. No. 1e-14;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MYLASSSTSIHTMLLMLFHLGLQASIS 30
 DB 1 MYLASSSTSIHTMLLMLFHLGLQASIS 30

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 19, 1999, 11:38:06 ; Search time 3012.52 Seconds
(without alignments) 12491.011 Million cell updates/sec

Title: US-09-037-657-38
Perfect score: 11832

Sequence: 1 gcgcgcgcgcgcgcgtgattac.....taagtctgggggtgggggga 11832

Scoring table: IDENTITY_NUC

Searched: 679419 segs, 1590154680 residues

Database :

GenEmbl.*
1: gb_bal.*
2: gb_bal.*
3: gb_bal.*
4: gb_bal.*
5: gb_bal.*
6: gb_bal.*
7: gb_bal.*
8: gb_bal.*
9: gb_bal.*
10: gb_bal.*
11: gb_bal.*
12: gb_bal.*
13: gb_bal.*
14: gb_bal.*
15: gb_bal.*
16: gb_bal.*
17: gb_bal.*
18: gb_bal.*
19: gb_bal.*
20: gb_bal.*
21: gb_bal.*
22: gb_bal.*
23: gb_bal.*
24: gb_bal.*
25: gb_bal.*
26: gb_bal.*
27: gb_bal.*
28: gb_bal.*
29: gb_bal.*
30: gb_bal.*
31: gb_bal.*
32: gb_bal.*
33: gb_bal.*
34: gb_bal.*
35: gb_bal.*
36: gb_bal.*
37: gb_bal.*
38: gb_bal.*
39: gb_bal.*
40: gb_bal.*
41: gb_bal.*
42: gb_bal.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	11830.4	100.0	11832	5	A70408	A70408 Sequence 38

2	6561.6	55.5	6663	5	A70398	A70398 Sequence 28
3	737.2	6.2	938	5	A70386	A70386 Sequence 16
4	737.2	6.2	1930	5	A70392	A70392 Sequence 22
5	607.4	5.1	40668	11	AC003112	AC003112 Human DNA
6	280.4	2.4	1629	5	A70382	A70382 Sequence 12
7	280.4	2.4	1673	5	A70384	A70384 Sequence 14
8	254.4	2.2	834	5	A70388	A70388 Sequence 18
9	220.2	1.9	1716	11	AF059293	AF059293 Homo sapi
10	190.8	1.6	1391	5	A70394	A70394 Sequence 24
11	141.2	1.2	560	5	A70393	A70393 Sequence 23
12	132	1.1	33613	12	AF060868	AF060868 Mus muscu
13	127.8	1.1	114898	12	AF139987	AF139987 Mus muscu
14	127.6	1.1	157684	12	AC004093	AC004093 Mus muscu
15	127.6	1.1	128026	12	AC004155	AC004155 Mus muscu
16	127.6	1.1	213245	12	MMHC425018	MMHC425018 Mus muscu
17	126.2	1.1	21984	12	MAA3128	MAA3128 Mus muscu
18	126	1.1	12300	12	MMBA3934	MMBA3934 Mus muscu
19	125.4	1.1	170351	34	MMBA3934	MMBA3934 Mus muscu
20	125.2	1.1	124230	12	AC005259	AC005259 Mus muscu
21	124.8	1.1	118235	12	AC003060	AC003060 Mus muscu
22	124.6	1.1	245439	12	MMHC135615	MMHC135615 Mus muscu
23	123.8	1.0	273800	12	MAA3CA212	MAA3CA212 Mus muscu
24	123.8	1.0	4784	12	MMCA313	MMCA313 Mus muscu
25	123.6	1.0	6089	12	MMCA313	MMCA313 Mus muscu
26	123.6	1.0	144723	12	AC005818	AC005818 Mus muscu
27	123.6	1.0	170965	35	AC005302	AC005302 Mus muscu
28	123	1.0	57327	12	AC003018	AC003018 Mus muscu
29	123	1.0	91638	34	AC002489	AC002489 Homo sapi
30	123	1.0	120479	35	AC006945	AC006945 Mus muscu
31	122.8	1.0	128143	12	AC003063	AC003063 Mus muscu
32	122.8	1.0	94459	12	MMTSDNA	MMTSDNA Mus muscu
33	122.8	1.0	128008	35	AC000996	AC000996 Mus muscu
34	122.6	1.0	140554	12	AC000399	AC000399 Genomic s
35	122.2	1.0	43576	12	AC003066	AC003066 Mouse cos
36	122	1.0	179436	35	AC006507	AC006507 Mus muscu
37	122	1.0	182721	35	AC005816	AC005816 Mus muscu
38	121.8	1.0	136687	12	AC005816	AC005816 Mus muscu
39	121.8	1.0	158405	12	MMHC310M6	MMHC310M6 Mus muscu
40	121.4	1.0	85931	12	AC007049	AC007049 Mus muscu
41	121.4	1.0	151218	35	AC005992	AC005992 Mus muscu
42	121.2	1.0	143	5	A70390	A70390 Sequence 20
43	121.2	1.0	29117	12	AC001230	AC001230 Genomic s
44	121.2	1.0	4101	12	MMCA312	MMCA312 Mus muscu
45	121	1.0	14318	12	AF136179	AF136179 Mus muscu

ALIGNMENTS

RESULT	1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES	BASE COUNT	ORIGIN
A70408	11832 bp	DNA	Sequence 38 from Patent WO9811225.	A70408	GI:4774683	unclassified.	unclassified.	unclassified.	1 (bases 1 to 11832)	Nicola, N.A., Fabri, L., Farley, A., Nash, A., Willson, T., Rakar, S., Zhang, J., Alexander, W., Hilton, D.J., Kojima, T., Medda, M. and Kikuchi, Y.	A NOVEL HAEMOPOLYMERIN RECEPTOR AND GENETIC SEQUENCES ENCODING SAME	NICOLA NIOS ANTONY (AU)	Location/Qualifiers	2447 a 3367 c 3298 g 2720 t	
A70408	11832 bp	DNA	Sequence 38 from Patent WO9811225.	A70408	GI:4774683	unclassified.	unclassified.	unclassified.	1 (bases 1 to 11832)	Nicola, N.A., Fabri, L., Farley, A., Nash, A., Willson, T., Rakar, S., Zhang, J., Alexander, W., Hilton, D.J., Kojima, T., Medda, M. and Kikuchi, Y.	A NOVEL HAEMOPOLYMERIN RECEPTOR AND GENETIC SEQUENCES ENCODING SAME	NICOLA NIOS ANTONY (AU)	Location/Qualifiers	2447 a 3367 c 3298 g 2720 t	

[illegible][illegible]

D	8641	GGGTTGGGGGGCTGGAGGACCACTAGATGTTGGCTGGGGTCTCACCAACAGTCTCAAG	8700
Q	8701	attctctcttccaaagccaaagtaaccagatccgcgtaccgcgttgaggacagcgttgcgtga	8760
D	8701	ATTCTCTCTTCCAAAGCCAAAGTACCAGATCCGATCCGATCCGGTGGAGAGACAGGTGACTGA	8760
Q	8761	aggtgcccgtttccggccccgggaacccgccttgaaacccgcgcgcgcgcgcgcgcgcgcgc	8820
D	8761	AGGTGCCCGTCTCCCGCCCCGGAGACCCGCCCTCGAACCCCGCCCCGGATCTGATCTCTCCCT	8820
Q	8821	caccgttcgaagtgtgtatgaatgaatcaaaccaacagacatctctgcgtctccgcggcgtga	8880
D	8821	CACCGTCGAGGTGTTGATGATACGTACAGCAACAGACCTCTCCGTCTCCGGGGCTTGA	8880
Q	8881	gcccggaaccgatttacttcgtccaaagtgcgtgtgaacccattcgggagctatlggtgcga	8940
D	8881	GCCGGGACCGTTACTTCTCTCCAAAGTGGCTTGAACCATTCGGAGATCTATGGGTGA	8940
Q	8941	aaagcggggaatctcgtggagcgaagtggggccaaccccaacgcgtcctcaacccctgaatgg	9000
D	8941	AAAGCGGGGATCTGGAGCGAGTGGAGCCACCACCCGCTGCTCACCCCTCGAATGG	9000
Q	9001	tgaacacctcccaaggcctgcgtcgtggccatggaatcccaatccatctcgttccctcccc	9060
D	9001	TGACACACTCTCCAGGGCTGCTGGCCCATGAGATCCCAATCCATCTGTTCTTCC	9060
Q	9061	cccaacccttttttgaagacagcgtctcttcagatgacgtgcgtgcgtcctaattcagatg	9120
D	9061	CCCAACCCTTTTGTGAGACAGCGTCTTCAAGTGGCCATGCGGCTTAATTCGATG	9120
Q	9121	tagtcaagatgaacctgcagctccgtgctcttttgcctccactgagacaatggcagct	9180
D	9121	TAGTCAGGAGTACCTGACACTCTGCTGCTTTTGTCTCACTTAGAGACATGGCCAGT	9180
Q	9181	ggccatcaccaaccccttggaggactagcatgtagcttattingcctctgcatttggtgaga	9240
D	9181	GGCCATCACCACTTTTGGGAGTACGCCATGAGTATTATTAAGCTGTCATTGGTGACA	9240
Q	9241	gatgagatacaacagttgtgacccctctgtlaagagaactgaagacagcgttcttaaccc	9300
D	9241	GATGAGATACAACAGTGTGACTCTTGTGAAGAACTGAAGACAGCGTTTTAACCC	9300
Q	9301	aatctcctaggtctctcctagaaggttaacctatataaataagagactattacagccagta	9360
D	9301	AATATCTTAGGCTCTCTAGAGGTTAATCTTATATAAAGAGACTATTACAGCCAGTTA	9360
Q	9361	tcaaatgtcccaacgaaccttttgcacaacaacatataagaccacagcgtcgtgctcac	9420
D	9361	TCAATGTCTCCACAGAACCTTTTGTACACAACCTATAGACCAAGTGCCTGTGCTTAC	9420
Q	9421	cacataagaggtctactatgtctggcccaacccctccaaaccttaaaaggttaacctagcagc	9480
D	9421	CACATAGAGGTCTTACTGTGTGGCCCAACCCCTTCAACCTTAAAGTAACTTAGCGACG	9480
Q	9481	cttaatatcttgcacatccctctactacccaagcctcttgaatgctcagaacacagcattaac	9540
D	9481	CTTATATTTGCAATCTCTCTACTACCTAGCCTCTTGATGATCTCAGAAACAGGCAATTAAC	9540
Q	9541	caaatctctctctctctggggccctttcttaaggttgggagggcctaagaatgaactcctctt	9600
D	9541	CAAGTCTCTCTCTCTGGTCCCTTTCTTAAGTGGAGGGGCTTAAGATGATCTTCTT	9600
Q	9601	gtctctgaagactctccagagcccaatgagatctgcacctctcaaatgaataatattgacata	9660
D	9601	GTCCTGAAGACTCTCCAGGCCCATGATCTGCACCTCTTAATATATAATATTGCAATTA	9660
Q	9661	aatgctcggcctcagtttccccacactgtcaggtttaggcagacagtcggtccaaagac	9720
D	9661	AATGCTGGCCTCAGTTTCCCCACACTGTCAGGTTTAGGACAGACAGTGGTCAAGACAC	9720
Q	9721	ttaattatttgcagcagtttaagaagaagctccatcccccaaccgcgttctctccgtgctc	9780
D	9721	TTCAATATTTCAGCGAGTATAAGAAGAACTCCATCCCCACCCCGCTCTCTCCGGTCC	9780

QY	9781	ctaaagacagaaataactcttcaactgtgaactgaaactcttcgcagacgcatatggtccacttaa	9840
Db	9781	ctTAaAaCaGAAATACTTCTTACACTGTGAACGAACCTCTCGAGACGGATATAGCTACACTTTAA	9840
QY	9841	tgatcatgaataaataaggggaaactgaagctccgagagatctcttgtaggaagaaggtcmaa	9900
Db	9841	TGATCATGAAATAAATAGGGGAACCTAGAGCTCCGAGAGATCTCTGGAGAGAAAGGGTCAA	9900
QY	9901	accagctccaaagaaactctcaagcccccaatccgggcctcccaagtlctgygcttgcgg	9960
Db	9901	ACCAGCTCCAGGAACCTCTCCAGCCCCCATCCGAGCTCTCCAGGTTCTGGGCTTGGCGG	9960
QY	9961	gaagtgaacaacaactctggggggggctctgaaagctctgagagacttggcccccttcgtggccagc	100200
Db	9961	GAGTGAACAACACTCTGGAGGGGCTTGAGGCTCTGAGACTTTGGCCCTTTGCTCTGGCCAC	100200
QY	10021	acctcgagatctcttgcaaggagcccaagcggctgcgclcgccccagagaagactgaagaag	100800
Db	10021	ACCTCGAGATTCTTGCAAGGGAGCCAGCAGGGCGGTGCGTCCGCCAGAGACTGAAGAAG	100800
QY	10081	ccggggggtagagttctgagaggagtgaaagagggctgctgtaggggcccgaagcttctgccagg	101400
Db	10081	CCGGGGGTAGAGTTCTGAGAGGAGGTGAAGAGGGGCTGTGGGGCCCAAGCTTGTGCCAGGG	101400
QY	10141	ccgtgcacagagctcccccagtttatattatgctgttaggcgcgaatgctctatccgtctggcc	102000
Db	10141	cCtGtCaCaGAGTtCCcCaGtTTtAtTtAtTtAGGcGTtAGcCcCaATtGtCtTtATtGcGTtGcC	102000
QY	10201	tgcctggggagatgctgcgcgcctgaggatcttgagcccaaggagctgccttcccaatagtcctc	102600
Db	10201	TGcTtGgGgGAtGcTgcGcGcCTtGgGAtTtGAGcCcCaAGGcGTtGcCTtCCaCTtCaCTtGcC	102600
QY	10261	caagccactccatgcatacaaccggctgcatctctctctgaggttactcttggaaacccggccctg	103200
Db	10261	CaGcCcAcTcCaTgcAtGtCaCaCcCcGtCaTtCTtGAGcCTtATtTtGgGaAcCcGcCcCTtG	103200
QY	10321	ttctgtgctgtctgtcctatcttccgtcaatcaacttcccaagagccttlttltatgctt	103800
Db	10321	TtCtGtGcTgTcTgTcTcTATtTtTtGtCaTtTtGtCaTtTtCCaAGcCTTtTtTtTtATtGcTt	103800
QY	10381	ttataataactacgttttaaaaatgctttgtataaagtgtgctgcctcgtgagcgtgc	104400
Db	10381	tTAAtAtTAACtAGtGtTTtAAAtAAAtGcTtTtTAAtAAAtGtGtGtCtTtCGtGAGcGtGc	104400
QY	10441	gtgcacaacaacaacgtgaaagtttagaagaacttggttgagtagagctctctccacaagt	105000
Db	10441	GtGcCaCaCaCaCaCaCgtGaaAGtTtAGAGAtTtGtTtGAGtAGAGcTcTtCCACaAtGt	105000
QY	10501	gggactagagctgcgcgaagaagacaattactgagtcactctgcgaagccccctcaaccctca	105600
Db	10501	GGGAcTAgAGcTgcGcGAGaAGAGcAAtTtACTGAGtCaTcTcCCAGcCCcCTCaCCcCTCA	105600
QY	10561	cttcccaatccgttggtgtagtgcataagtaatcgagaagtaatcgctgagcttaatttcg	106200
Db	10561	cTtCCcAAtCcGtTtGtGAtAGtAGtAAGtAAAtCGAGGtAAAtCcTtGcCTTtAAAtTtGc	106200
QY	10621	tagcatctcgtcctagaactcaacaagtgcgtgctcaccagcttctgtagagggagctctcc	106800
Db	10621	TAGcATcTcGTcGcTcAGcCTtACCaAGtGcTGTGtTtACCaCGtTtGTGGAGGGGcTcTCC	106800
QY	10681	ttccaggtctctgtaggggtatacaagctcccaagatctctgcttctctctctctctctctctct	107400
Db	10681	tTCCAGGtTcTtGgGgGtAtaCaAGtCCCaAGAtCtTcTtGCTTtCTTtAGGtCTTtGTCTtTAAGT	107400
QY	10741	tgcacctgcttgcgcgtgtccctcagaagctcccgcccaactgagcttccatctatctc	108000
Db	10741	TGcCCcTtGcTtTtGtGcGtGtCCtTAAGtAGtAGtAAAtCGAGGtAAAtCcTtGcCTTtAAAtTtGc	108000
QY	10801	ctttctgcagaaatacttcggttttaactctccacatgatttgacccctccctcttgcgtgct	108600
Db	10801	cTtTtCTtGcAGAAAtcTcGtGtTtTtACtTtCCACAGtAGtTtGAtCCcTtCTTtGcTtGtTcT	108600

QY	5903	gttaacagcttaagtccttgatcaatcaactaaggtctcgaagcttagccacagctctctt	5902
Db	732	gtttcacagttatgctcctgattcaacacpaaaggtcctcaggttgccacagctctctt	791
QY	5963	atgacacgagccaattcaacgagcccgatgttaccagatgaataatttttgaataagc	6022
Db	792	atgtactagaccatttcactcgtgccccttggaattgactgaatgaatttttttgatpaaagt	851
QY	6023	ctcttgaagctcctagcttaggctcaaaactaagtaccccaaggttcactcttgagctgct	6082
Db	852	ctctgttagcttagagcttaggctcaaacctpaaactgaaactgccaaaggtcatcttgagctgctgct	911
QY	6083	acctgtcttccaccccacaagtggtgatgatatactcaaggcagacctctctcggggaagg	6142
Db	912	acttttctctccaccccacaagtggtgatgatatactcaaggcagacctctctcggggaagg	971
QY	6143	gcttgcccttgagcttgaatttgttgctgaattcaatgaagcttggtctcgttctt	6202
Db	972	gcttgcccttgagcttgaatttgttgctgaattcaatgaagcttggtctcgttctt	1031
QY	6203	ctttctcttaactcgttgaatggttgagacacccgttcaaaagcttccctgactcttgaacat	6282
Db	1032	ctttctcttaactcgttgaatggttgagacacccgttcaaaagcttccctgactcttgaacat	1091
QY	6263	ccagagcaggttgagggagcttgaagtgtggtctatcccatgctctacaagaatgtcgtcttg	6322
Db	1092	ccagagcaggttgagggagcttgaagtgtggtctatcccatgctctacaagaatgtcgtcttg	1151
QY	6323	accccaacacagctgtaactcagccccccaggaacccacacctctcatcgtctctccctg	6382
Db	1152	accccaacacagctgtaactcagccccccaggaacccacacctctctcatcgtctctccctg	1211
QY	6383	caagctacccgtcctatatactatgagacacacacacctcgggccaacgcctcgaagggctctacgg	6442
Db	1212	caagctacccgtcctatatactatgagacacacacacctcgggccaacgcctcgaagggctctacgg	1271
QY	6443	accctcaatgtgcgcgcgcctgcctctgagctgtgccgcctctcttaaacacctccacccctg	6502
Db	1272	accctcaatgtgcgcgcgcctgcctctgagctgtgccgcctctcttaaacacctccacccctg	1331
QY	6503	gaccttgacctgtgctaaacttaatgggtccagcagcatcagggagacaatctggtgtgt	6562
Db	1332	gaccttgacctgtgctaaacttaatgggtccagcagcatcagggagacaatctggtgtgt	1391
QY	6563	caagcccgagacgcaagcatctctgctgtgcctcctgtcctatgcttggctggaatggtggc	6622
Db	1392	caagcccgagacgcaagcatctctgctgtgcctcctgtcctatgcttggctggaatggtggc	1451
QY	6623	cccaagacactcaagatagatagatgggggttggaatgacaaatttagagcttggtcttg	6682
Db	1452	cccaagacactcaagatagatagatgggggttggaatgacaaatttagagcttggtcttg	1511
QY	6683	tctctgggagagacgaatgggtctcactcgtcatgagcgaatggtcataccagcaagc	6742
Db	1512	tctctgggagagacgaatgggtctcactcgtcatgagcgaatggtcataccagcaagc	1571
QY	6743	atgccaactcaaggaagactgtggtgtgacacgttcccccgtgttaacccaagaactttaa	6802
Db	1572	atgccaactcaaggaagactgtggtgtgacacgttcccccgtgtgttaacccaagaactttaa	1631
QY	6803	aagctgtcatgtttctctcttgtagtgcaccttgaaagccctttaacatgaagctgtgtc	6862
Db	1632	aagctgtcatgtttctctcttgtagtgcaccttgaaagccctttaacatgaagctgtgtc	1691
QY	6863	ccggaaacatgaagatctcaacgtgcgccttgaaacccgggtgcacacgggggaacacttctt	6922
Db	1692	ccggaaacatgaagatctcaacgtgcgccttgaaacccgggtgcacacgggggaacacttctt	1751
QY	6923	aataccaacactccctcaagatnacaaagctggaagttgtgaaacagcgcaagcctgtgtg	6982
Db	1752	aataccaacactccctcaagatnacaaagctggaagttgtgaaacagcgcaagcctgtgtg	1811

QY	6983	tgactcttggaactacttactctctctgtatcaaatstgctctgttttltgaactcaaaag	7042
Db	1812	tgactcttggaactacttactctctgtatcaaatstgctctgttttltgaactcaaaag	18711
QY	7043	ggaactctgcacactccacaggttggtacaggttcagataacacatgtgagagatcacacact	71020
Db	1872	ggactcttcgcacctccacaggttggtacaggttcagataacacatgtgagagatcacacact	19311
QY	7103	gtgggacctactaatgtccatatcccaaggaaactgtgcctcttcaactccatgtagatc	7162
Db	1932	gtgggacctactaatgtccatatcccaaggaaactgtgcctcttcaactccatgtagatc	1991
QY	7153	tgagttgaaagccaccaaactgcctcctaagctcaagacagatctcgatgtccctcaactgtgac	7222
Db	1992	tgagttgaaagccaccaaactgcctcctaagctcaagacagatctcgatgtccctcaactgtgac	2051
QY	7223	ctgagagctgggtgagcccccaagtgctcacactgtgtctctccttagaaccttaaggagccc	7282
Db	2052	ctgagagctgggtgagcccccaagtgctcacactgtgtctctccttagaaccttaaggagccc	21111
QY	7283	tccccccatcccccccaagacttttggctctcttagaggctcttaagcacaagcagcgtgag	7342
Db	2112	tccccccatcccccccaagacttttggctctcttagaggctctcttaagcacaagcagcgtgag	2171
QY	7343	ctgcagaagacagtggtgtgtctcaactaatgtcacaaagactttcccccaagacgttcaaga	7402
Db	2172	ctgcagaagacagtggtgtgtctcaactaatgtcacaaagactttcccccaagacgttcaaga	2231
QY	7402	ctttccacctcccccccaacacaaabatacacaacactcttgaaagaaacactgtgac	7461
Db	2232	ctttccacctcccccccaacacaaabatacacaacactcttgaaagaaacactgtgac	2291
QY	7462	tgaccacctctccctctcttaagaccagagtggttcaagaaggagctccctaaaggagactgaagag	7521
Db	2292	tgaccacctctccctctcttaagaccagagtggttcaagaaggagctccctaaaggagactgaagag	2351
QY	7522	agagcgccccaggtctgtaaagcgcccccaaggaaagccagagccttgagctgtggggggggcgga	7581
Db	2352	agagcgccccaggtctgtaaagcgcccccaaggaaagccagagccttgagctgtggggggggcgga	2411
QY	7582	gggtttgagagagcaactgtgatatccctbagaacaaactggggccctaatctaatcttagtggtg	7641
Db	2412	gggtttgagagagcaactgtgatatccctbagaacaaactggggccctaatctaatcttagtggtg	2471
QY	7642	ttccccagcccaaaagcagcgcggggccatttaacacctcaagtccctcaactgaagactcagag	7701
Db	2472	ttccccagcccaaaagcagcgcggggccatttaacacctcaagtccctcaactgaagactcagag	2531
QY	7702	ggagagatcagctgtgtactctctctccatgtgtcccccagagaggttccctgtgtccccctggc	7761
Db	2532	ggagagatcagctgtgtactctctctccatgtgtcccccagagagaggttccctgtgtccccctggc	2591
QY	7762	tcatttcccaacatcccaagaggtttgtgtgtcttccctcgagatcaacctcagttgtctcgt	7821
Db	2592	tcatttcccaacatcccaagaggtttgtgtgtcttccctcgagatcaacctcagttgtctcgt	2651
QY	7822	ggcttgacacagctgcctcccgctggagagcctctgtgtatctgtacaagagcactcagaggtgagat	7881
Db	2652	ggcttgacacagctgcctcccgctggagagcctctgtgtatctgtacaagagcactcagaggtgagat	2711
QY	7882	ggagatvgggatatacagaaggtgagccaaatagcacctccaaggtgggggtgtatatacaata	7941
Db	2712	ggagatvgggatatacagaaggtgagccaaatagcacctccaaggtgggggtgtatatacaata	2771
QY	7942	aagcttgtcaacctcgaagcttcagaaagcctactccatcgatgtatcaacaatgtgtgacataac	8001
Db	2772	aagcttgtcaacctcgaagcttcagaaagcctactccatcgatgtatcaacaatgtgtgacataac	2831
QY	8002	tcctggagacatgttagtgagaccctcagcttcaaaaacagacagcttagctttaagagttcagctt	8061
Db	2832	tcctggagacatgttagtgagaccctcagcttcaaaaacagacagcttagctttaagagttcagctt	2891
QY	8062	gtgacttaatactctgaactcgaagccttaataggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	8121

```
|||||
Db 2892 GAGACTTAATGAGAACTGAGGCGCTATAGTGTGTGGTGTATGCTCCGCTACTCTCCCT 2951
QY 8122 GTTGTGTAAGATCTCTGCGCTAACTCCAGCCAGCTGGTGGTCTGCTCTGCTCCCTG 8181
Db 2952 GTTGTGTAAGATCTCTGCGCTAACTCCAGCCAGCTGGTGGTCTGCTCTGCTCCCTG 3011
QY 8182 AGGGAGGAATGTGTGTTTCCATCAAGAGATAGGACCCCTGTAGACGCAACTGCTGCTG 8241
Db 3012 AGGGAGGAATGTGTGTTTCCATCAAGAGATAGGACCCCTGTAGACGCAACTGCTGCTG 3071
QY 8242 GCTGTTTCTGGAATTTAAATGACGATATCTATAGGCTTGAGAGTAACTAAGG 8301
Db 3072 GCTGTTTCTGGAATTTAAATGACGATATCTATAGGCTTGAGAGTAACTAAGG 3131
QY 8302 GTGGGGGCGTGTCTGTAAGAAAGCAGATAGGCTATAGGACCATGCACTGAATATTC 8361
Db 3132 GTGGGGGCGTGTCTGTAAGAAAGCAGATAGGCTATAGGACCATGCACTGAATATTC 3191
QY 8362 ACCAGTGGTGTCTGTCTACAGGCAATCTCAACAGCAAGTCTCAAGCACTGGAGCAC 8421
Db 3192 ACCAGTGGTGTGTCTGTCTACAGGCAATCTCAACAGCAAGTCTCAAGCACTGGAGCAC 3251
QY 8422 TGTGCGACGATTTAAATGACGATTTAAATGACGATTTAAATGAGGAGGAGGAGGAGG 8481
Db 3252 TGTGCGACGATTTAAATGACGATTTAAATGACGATTTAAATGAGGAGGAGGAGGAGG 3311
QY 8482 ATCTCTGTAGTCAAGGCACTCTGAATTTAAATGAGGACTCCAGGCCAGGAGGAGG 8541
Db 3312 ATCTCTGTAGTCAAGGCACTCTGAATTTAAATGAGGACTCCAGGCCAGGAGGAGG 3371
QY 8542 CGGAGTAAACCTGTCTCAAAAAAAGCAATCTTGTAGTGTGTGTGTGTGTGTGTGTGTGT 8601
Db 3372 CGGAGTAAACCTGTCTCAAAAAAAGCAATCTTGTAGTGTGTGTGTGTGTGTGTGTGTGT 3431
QY 8602 CAGTGAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 8661
Db 3432 CAGTGAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3491
QY 8662 AGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 8721
Db 3492 AGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3551
QY 8722 ACCAGATCGCTACCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 8781
Db 3552 ACCAGATCGCTACCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3611
QY 8782 CCGGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 8841
Db 3612 CCGGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3671
QY 8842 CCGGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 8901
Db 3672 CCGGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3731
QY 8902 CCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 8961
Db 3732 CCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3791
QY 8962 GTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 9021
Db 3792 GTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3851
QY 9022 GTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 9081
Db 3852 GTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3911
QY 9082 GTGTCTCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 9141
Db 3912 GTGTCTCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3971
QY 9142 TCTGTGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 9201
|||||

Db 3972 TCTGTGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4031
QY 9202 ACTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 9261
Db 4032 ACTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4091
QY 9262 CTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 9321
Db 4092 CTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4151
QY 9322 GTTAACCTTATTAATAAATGAGAGATATTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 9381
Db 4152 GTTAACCTTATTAATAAATGAGAGATATTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4211
QY 9382 TTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 9441
Db 4212 TTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4271
QY 9442 GTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 9501
Db 4272 GTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4331
QY 9502 AGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 9561
Db 4332 AGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4391
QY 9562 CTTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 9621
Db 4392 CTTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4451
QY 9622 CATGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 9681
Db 4452 CATGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4511
QY 9682 CACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 9741
Db 4512 CACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4571
QY 9742 AAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 9801
Db 4572 AAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4631
QY 9802 CTGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 9861
Db 4632 CTGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4691
QY 9862 ACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 9921
Db 4692 ACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4751
QY 9922 AGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 9981
Db 4752 AGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4811
QY 9982 GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10041
Db 4812 GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4871
QY 10042 GTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 10101
Db 4872 GTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4931
QY 10102 GTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 10161
Db 4932 GTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4991
QY 10162 GTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10221
Db 4992 GTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5051
QY 10222 GTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 10281
Db 5052 GTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5111
```


[illegible]

LOCUS AC003112 40668 bp DNA PRI 21-NOV-1997
 DEFINITION Human DNA from chromosome 19 specific cosmid R30292, genomic
 sequence, complete sequence.
 ACCESSION AC003112
 NID 92636669
 VERSION AC003112.1 GI:2636669
 KEYWORDS HTG
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 40668)
 Lamerdin, J.E., McCreedy, P.M., Adamson, A.W., Burkhart-Schultz, K.,
 Gordon, L., Christensen, M., Kyle, A., Ramirez, M., Stillwagen, S.,
 Garms, J., Dangnan, L., Bruce, R., Quan, G., Montgomery, M., Ov, D.,
 Kobayashi, A., Olsen, A.O. and Carrano, A.V.
 Sequence analysis of an ~1 Mb region containing the MER2B gene in
 19p12
 19p12
 TITLE Unpublished
 JOURNAL 2 (bases 1 to 40668)
 REFERENCE Lamerdin, J.E.
 AUTHORS Direct Submission
 JOURNAL Submitted (21-NOV-1997) Human Genome Center, Lawrence Livermore
 National Laboratory, 7000 East Ave., Livermore, CA 94551, USA
 FEATURES
 source
 1.40668
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="R30292"
 /chromosome="19"
 /map="19p12 between UBA52 and D19S451"
 /cell_line="5H2-B"
 /clone_lib="LL19NC03 R chromosome 19 cosmid library"
 /note="LL19NC03 cosmid library constructed at LLNL from
 flow-sorted chromosomes from hybrid 5H2-B, which carries
 chromosome 19 as its only human chromosome."
 complement(6..62)
 repeat_region
 msc_feature
 /rpt_family="Alu"
 /note="predicted exon, program: grail2exons_human_1.3,
 frame: 1, quality: good, score: 71.000"
 complement(1383..1444)
 msc_feature
 /note="predicted exon, program: grail2exons_human_1.3,
 frame: 1, quality: good, score: 62.000"
 2257..8516
 /standard_name="endogenous retroviral sequence"
 /note="HERV9 retroviral sequence"
 complement(3269..3544)
 repeat_region
 /rpt_family="Alu"
 3425..5722
 /rpt_family="Alu"
 complement(6486..6772)
 repeat_region
 /rpt_family="Alu"
 complement(7505..7783)
 repeat_region
 /rpt_family="Alu"
 8931..9254
 /rpt_family="Alu"
 9358..9535
 /rpt_family="L1R12"
 complement(9445..9505)
 /note="predicted exon, program: grail2exons_human_1.3,
 frame: 1, quality: good, score: 63.000"
 msc_feature
 /note="predicted exon, program: grail2exons_human_1.3,
 frame: 1, quality: excellent, score: 75.000"
 complement(9928..10123)
 repeat_region
 /rpt_family="Alu"
 10409..10548
 /note="DSS similarity to AA047548 z15e02.r1 Soares fetal
 heart NBH19W Homo sapiens cDNA clone 377018 5' (1..138);
 96% identity.--Other overlapping matches:--(10435..10548)
 DSS similarity to AA136115 zK90B04.r1 Soares pregnant
 uterus NBHPU Homo sapiens cDNA clone 490063 5' (1..110);

93% identity.--(10466..10548) DSS similarity to AA452628
 zX33f04.r1 Soares total fetus NB2HF8 9W Homo sapiens cDNA
 clone 788287 5' (1..82); 95% identity.--(10486..10548) DSS
 similarity to AA094412 z82h02.r1 Soares fetal heart
 NBH19W Homo sapiens cDNA clone 365523 5' (1..61); 97%
 identity."
 10643..10725
 /note="predicted exon, program: grail2exons_human_1.3,
 frame: 1, quality: excellent, score: 100.000--DSS
 similarity to AA047548 z15e02.r1 Soares fetal heart
 NBH19W Homo sapiens cDNA clone 377018 5' (139..221); 100%
 identity.--DSS similarity to AA136115 zK90B04.r1 Soares
 pregnant uterus NBHPU Homo sapiens cDNA clone 490063 5'
 (111..192); 99% identity.--DSS similarity to AA452628
 zX33f04.r1 Soares total fetus NB2HF8 9W Homo sapiens cDNA
 clone 788287 5' (83..165); 100% identity.--DSS similarity
 to AA009412 z82h02.r1 Soares fetal heart NBH19W Homo
 sapiens cDNA clone 365523 5' (62..144); 100% identity."
 11127..11331
 /note="DSS similarity to AA136115 zK90B04.r1 Soares
 pregnant uterus NBHPU Homo sapiens cDNA clone 490063 5'
 (193..394); 94% identity.--DSS similarity to AA094412
 z82h02.r1 Soares fetal heart NBH19W Homo sapiens cDNA
 clone 365523 5' (145..386); 96% identity.--(11121..11331)
 DSS similarity to AA09693 z82h02.s1 Soares fetal heart
 NBH19W Homo sapiens cDNA clone 365523 3' (429..307); 95%
 identity.--(11127..11246) DSS similarity to AA047548
 z15e02.r1 Soares fetal heart NBH19W Homo sapiens cDNA
 clone 377018 5' (222..341); 97% identity.--(11061..11331)
 predicted exon, program: grail2exons_human_1.3, frame: 0,
 quality: good, score: 73.000--(11061..11209) DSS
 similarity to AA45628 zX33f04.r1 Soares total fetus
 NB2HF8 9W Homo sapiens cDNA clone 788287 5' (167..313);
 98% identity."
 complement(11869..12161)
 /rpt_family="Alu"
 12502..12581
 /note="DSS similarity to AA136115 zK90B04.r1 Soares
 pregnant uterus NBHPU Homo sapiens cDNA clone 490063 5'
 (395..477); 92% identity.--DSS similarity to AA094412
 z82h02.r1 Soares fetal heart NBH19W Homo sapiens cDNA
 clone 365523 5' (347..427); 94% identity.--DSS similarity
 to AA09693 z82h02.s1 Soares fetal heart NBH19W Homo
 sapiens cDNA clone 365523 3' (306..228); 99%
 identity.--DSS similarity to AA450010 zX33f04.s1 Soares
 total fetus NB2HF8 9W Homo sapiens cDNA clone 788287 3'
 (309..229); 99% identity."
 complement(12966..13740)
 /rpt_family="Alu"
 13740..13826
 /rpt_family="Alu"
 complement(13274..13531)
 /rpt_family="Alu"
 complement(13756..14029)
 /note="predicted exon, program: grail2exons_human_1.3,
 frame: 2, quality: good, score: 61.000--Other overlapping
 matches:--(13756..13826) DSS similarity to AA094412
 z82h02.r1 Soares fetal heart NBH19W Homo sapiens cDNA
 clone 365523 5' (428..496); 97% identity.--(13756..13984)
 DSS similarity to AA09693 z82h02.s1 Soares fetal heart
 NBH19W Homo sapiens cDNA clone 365523 3' (227..11); 98%
 identity.--(13756..13984) DSS similarity to AA450010
 zX33f04.s1 Soares total fetus NB2HF8 9W Homo sapiens cDNA
 clone 788287 3' (228..11); 99% identity."
 complement(14100..14672)
 /rpt_family="Alu"
 14685..15317
 /note="DSS similarity to AA406406 zV11e07.s1 Soares NBHPU
 SI Homo sapiens cDNA clone 753348 3' (1..433); Score: 858
 identity: 431/433 (99%)--(14684..15237) DSS similarity
 to K37175 zB21e02.r1 Soares fetal lung NBH19W Homo
 sapiens cDNA clone 302666 5' (1..355); 94%
 identity.--(15227..14885) DSS similarity to AA121532
 zK89c11.s1 Soares pregnant uterus NBHPU Homo sapiens cDNA
 clone 490004 3' (342..11); 99% identity.--(15227..14885)

[illegible]

OY	1226	gagcgccacacttggaactcctccgggaacaagaaggcg-----gccttcgggg	1274
Db	28108	GGGGCGCCACCTTGAGGACCCTGGGAACAAGAAGGCCCCCCGGCGCAGCCCTTCGGGG	28049
OY	1275	cgcctcaacctgttgggtcatalcgtgcaccaaccacagcctcccaaagtaccggtata	1334
Db	28048	CGCCTTCAACTGTGGGGGACACAGGGGGCGCATTCACACTTCCAGAGAGTCTGTCCGC	27985
OY	1335	catcaaggcctctatctgtatcccctttgcgaagctgtctgcgaagctcaattgaa	1394
Db	27988	GCCCAGGCCCTGTAACTGGGGCACATGGGCGGAGGGGSCGAGGCCAGGCCCCCGGGGGC	27929
OY	1395	ggaactcgaaagtctcttggagacccccctctcaagtgtagtgtaagcttggggcac	1454
Db	27928	TGGGGGAAATCGGGGCGCTTGAGACCTCTCGATCCAGAGAGTAGGGGGCCACGCGTTGC	27869
OY	1455	gctctgtcttgyatatcagaagcggaaagaaagctcctctgcgggggcgcagcttg	1514
Db	27868	TTCCGTGGAATGTGGGCCCTG-GGACTGTGAAGGACCGCTCCAGGCGGGGGCCAGCG	27810
OY	1515	gtcgcttgtgttgggtgtcgtgcgcaaatgttgggtccccccccaatyaagtatccc	1574
Db	27809	GATTGGCCCGCTGGGTGCGGGCCCCACAGCGG---CCCCCCAGGGGAGGATGAGCCC	27533
OY	1575	cgggggggggtgtggggcggtatcgyagcgccctctgcgtcgtgtgatcggccgggt	1634
Db	27732	CACGGGGAGGGGGGGG-C-TGCAGGGGGGGCGCGCTGTGGCGCTTGCGCCACCGGGGT	27694
OY	1635	ccctcggaactgtcctctccgtygggtgtcgcgcgccccctccccctatagcaactcc	1694
Db	27693	CCCCGGGGAATTGTGTCTTCCACGGGCGCGGCCACCGCCTCCCC--TGGGGGCCCG	27637
OY	1695	atgcttgytatcctcgaagctcctcctcactgtgttgggtctcaaacggtctcatcaag	1754
Db	27636	GGCGACGGGGGGTCCCGGAGTCCCGGCACCGCGGGCGGTGAACCGTTTTCAOCGAG	27577
OY	1755	ctgcgtgtgttgaagaagcctaagcgaactgaattcgtga-----	1797
Db	27576	CTGCTCTCGGGCGAAGCGCTCCAGCGAGGAGTAATTGTGTGAGCAGTAGTGAGAGGGA	27517
OY	1797	---ggaaggaagaaagcgtgtccgggcaaccggccccagctcatgtctaagggtc	1853
Db	27516	GAGGAGCGGGGAGCCAGTGTAGTCCGGGCKCCAGCGGGCCAA---ACTTAGGGGGGC	27461
OY	1854	aaccagttgggtgcagctgcgcgagaaatcccaactgtcccaagaggaactctgcgttg	1913
Db	27460	CACGTGGAAACCGGTGTCCCTTAAGGGGCCAAGAACCTCTGCCCTGTAGAGGACTTGTGCCCTA	27401
OY	1914	agcccc-----ataccccaagcaacatcccgccagatgtcgtctcaatccaga	1968
Db	27400	GGCACTGTGGGGCCCCCATCCAAAGAACACCCCGCAAGATAGCTTCCGTATGACGA	27341
OY	1969	ccctcttgggaacacaccacaagaacacaaaaggccccactgtgtatgtgccgtcac	2028
Db	27340	CCCGCTTGGGACACACCCAGAGAAC-AAGATATCAACTGTGCTGTGGCCGCTCAC	27282
OY	2029	cctgcctccga-----cgcgctgcgcagccagatgcgtattgcacaacatcg---	2081
Db	27281	CCTGCCCTCAAACCTGGGCACTGTGCATCTCCAGTGTGCATTTCCGGCACCTTCGGGAC	27222
OY	2081	gcgtcgcatctcatcctctacacacacacacacacacacacacacacacacacac	2140
Db	27221	GCACACATAATTCATCCCGGACACACACACTTACAGGTACACATATCC-----	27173
OY	2141	acacacagaacgcac	2200
Db	27173	--CACAGGGGAGCGACCCCAAGGGGCACAGAGCGACCACTGCCCCAGATTTTCAG	27117
OY	2201	acattatttcacaggggggcaacacgggttacgcatagtgttagtgcactgtgagat	2260
Db	27116	TCACCTCACGAC-----	27103

QY	2261	cttcccccacactctcagaccaccatccatccgagacacagccacacgcgaagggacac	2320
Db	27103	----CCATCCAGCGGCACGCGGCCACCCCGGCGTCCGGCCGACACACTGCTCCAGGC	27048
QY	2331	gtcgcgcgtcgtctctcgtggcctagtagtctgtg-aaattgtccgagtgatctgtgaag	2379
Db	27047	TCATCCACCCCTGTGTGGCTGTGTCGGGGCTGCCTGTAGGGCTCTCCCGGAAAGGGCA	26988
QY	2380	ccctccgcctctgtcaggggacagaaacctacactcctgtctgtgccaaaggggtctggc	2439
Db	26987	GGGTCCGGGCATCTTCTTAGGGGCACTTCTGAAGACCTGGCTCTCCCTGGCTTTCAGT	26928
QY	2440	aggtatgtgtgtacaccccgagacccttcgcggagatgtgtgtctgtcgaagccttgg-	2499
Db	26937	GAG-----GGGCTACCTCGGGCCCTCCCAAGGCCCCCTGACGGCTCTCCAGGCGGTGG	26874
QY	2499	--tagttttaatgcccacaatagsgttaaagcttgttccggcggtctgcgaagcaa	2556
Db	26873	CAGCCCTGTAGTGCCTACCTGAGCGGCAAGCTTTGTT---TCACCCAGTAAGCAG	26818
QY	2557	cagcgaaagtggtcggaatgggggtggcgc-----gtgtgttcttctttaaaggggag	2611
Db	26817	CAGGCAAGGTGTAGGGGGTGGGGGGCACTAGTGTGTTTTTCTTTAAGGGGGA	26758
QY	2612	agaatttaataagsggtttcccaacccctcgcactcgttgttaactaccgtgttctta	2671
Db	26757	AGAAATTAATAAAGATTCTACACCTCGGCAAAATTTCTACTTACGGGTGTCCTTA	26698
QY	2672	aacactgcacacccagccggtgtgtgtcgttaaaagttaacagatcacagcggaacagg	2731
Db	26697	GCACTGACAGCAGCGCGGGTGGTGGT-AAAAGTCTGCAAGTACAGCGGAGGAGGAG	26659
QY	2732	atgggggcccctgtggtatgtgcgtggatlgagagccactccocgttgccttccaggaa	2791
Db	26638	ATGGGACCCCTGTGGGAGCCCTGGGAAGGAGGCCACCTTCCGCCCTGTGCAGAGAA	26579
QY	2792	tctcacacttccctttaaacaacatggtgtt-ctttaaataagcgagcaactcgc	2850
Db	26578	TCTCACTTTTCCCTTTTAAACACATGtGTtTTCITTTTAATACGGCAGTGGCTCCA	26519
QY	2851	catgtggaaggggga--aaataagctgtatagccccgcttgttggaaagagggga	2907
Db	26518	GATTGGGAAAAGAGAGAGAAAAAATCTTTAAAGCCCCAGCTGTGGGGAAGAGAGGG	26459
QY	2908	ag--agggaaagaaaaagagaggtgtctcctcaagctttagggggcgc	2954
Db	26458	AGCGAGGAGTGAAGGAGACTGTTCAAAAGCCCACTACATGAGGATCTG	26410

RESULT	6
LOCUS	A70382
DEFINITION	A70382 1629 bp DNA
ACCESSION	A70382
NID	94774663
VERSION	A70382.1
KEYWORDS	GI:4774663
SOURCE	unidentified.
ORGANISM	unidentified
REFERENCE	unclassified.
AUTHORS	1 (bases 1 to 1629)
	Nicola,N.A., Fabril,L., Farley,A., Nash,A., Willson,T., Rakar,S.,
	Zhang,J., Alexander,W., Hilcon,D.J., Kojima,T., Maeda,M. and
	Kikuchi,Y.
TITLE	A NOVEL HAEKPOIETIN RECEPTOR AND GENETIC SEQUENCES ENCODING SAME
JOURNAL	Patent: WO 9811225-A 19-MAR-1998;
	NICOLA NICOS ANTONY (AU)
FEATURES	location/Qualifiers
source	1..1629
	/organism="unidentified"
	/db_xref="taxon:32644"
BASE COUNT	336 a 341 c 453 g 299 t
ORIGIN	

Db	354	ACGGGGCCCGCTCCCGCCCGCTCCAGCTCTCCCTGTACCTCAACAGCCTCCACCTGGCTCGG	413
QY	6510	cccttgctaaccttaattatggtctccaggagcagtlcaggagacaacatctggtgtcagccc	6569
Db	414	CCCTGGCCCACTCAATGAGGTGTCAGGACGAGCGGTGGGGGACCACTCGTGTGCCACGCC	473
QY	6570	gagacggcagcagctcttgctggtcctcgtcctctatgtgtgc	6610
Db	474	GTGACGGCAGCATCTGCTGGCTGCTCCTCATATGTTGG	514
RESULT	10		
LOCUS	A70394	1391 bp	DNA
DEFINITION	Sequence 24 from Patent WO9811225.		PAT
ACCESSION	A70394		07-MAY-1999
NID	94774672		
VERSION	A70394.1	GI:4774672	
KEYWORDS			
SOURCE	unidentified.		
ORGANISM	unidentified.		
REFERENCE	1 (bases 1 to 1391)		
AUTHORS	Nicola, N.A., Fabril, L., Farley, A., Nash, A., Willson, T., Baker, S., Zhang, J., Alexander, W., Hilton, D.J., Kojima, T., Maeda, M., and Kikuchi, Y.		
TITLE	A NOVEL HAEMOPOIETIN RECEPTOR AND GENETIC SEQUENCES ENCODING SAME		
JOURNAL	Patent: WO 9811225-A 19-MAR-1998;		
	NICOLA NICOS ANTONY (AU)		
FEATURES	Location/Qualifiers		
source	1..1391		
	/organism="unidentified"		
	/db_xref="taxon:32644"		
CDS	<1..1053		
	/note="unnamed protein product"		
	/codon_start=1		
	/protein_id="CAB42576.1"		
	/db_xref="PID:e1433703"		
	/db_xref="PID:g474673"		
	/db_xref="GI:4774673"		
	/translation="TLNGRLPELPSLVNASTLALANLNGSRSGDNVICHARD GSLIAGSCLVYGPREPEPVNISGMSKMKMLTQRMPGAGGEFLHNYSLAKYLRKY GQNTCEEHHTVQPHSCHITKDALFLPEIWEATKRLSANSVDLILDLVYITD PPDVAHSRAGLELDLSVWVSPALKDLFQAKQIRRVSDVDMKVVDVDSNQT SCLAGLKPTGVYFVQRCNPFGIYGSKKAGINSEWSHPAASTRPSERGPGGACE PRGEPSSGCVREELKQFLGMLKKHAYCSNLSPRLYQWBRAMWQSKHTNRHRTSS CPRADGRREVLDKX"		
BASE COUNT	281 a	459 c	417 g
ORIGIN			234 t
Query Match	1.6%; Score 190.8; DB 5; Length 1391;		
Best Local Similarity	73.5%; Pred. No. 2.5e-30;		
Matches 291; Conservative 0; Mismatches 37; Indels 68; Gaps 1;			
QY	8604	gtaccacgagaccaccaccacgacgltgcacgltgacgcggttggggcgcttggaagaccag	8663
Db	469	GTGACGACGAGACCCCGCCCGACGTGCAGCTGAGCCGCGCGGGGCGCTGAGAGACCAG	528
QY	8664	ctgaggttggtggtctcaccacacgcttcaagatctctctccaagccaagtac	8723
Db	529	CTGAGCGTGGCCTTGGTGTCCACCCGCTCAAGGATTTCTCTTCAAGCCAAATAC	588
QY	8724	cagatcgctacgcgctgggaagacagcgttgacgcyggaagtgcccgctcccgccgagcc	8783
Db	589	CAGATCCGCTACCAAGTGGAGACAGTGTGACTGGA-----	626
QY	8784	cgcccttgaccgcccccgcgcacatctactctcctccacacgltgcaggtgtgatgacg	8843
Db	626	-----AGGTGTGACGAGT 640	
QY	8844	tcagcaaccagacctctctgcgcttgcgcggcgtaagaccgagcaccglttaactgtctc	8903

RESULT	13		
LOCUS	AF139987/C		
DEFINITION	AF139987	114898 bp	DNA
ACCESION	AF139987		03-JUN-1999
VERSION	94572848		
KEYWORDS	AF139987.1	GI:4972948	
SOURCE	house mouse.		


```

RESULT 14
AC004093/c AC004093 157684 bp DNA ROD 31-JAN-1998
LOCUS Mus musculus (129SV) DNA, unmapped BAC 10817, complete sequence.
DEFINITION AC004093
ACCESSION 92828320
VERSION AC004093.1 GI:2828320
KEYWORDS Htg.
SOURCE house mouse.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 157684)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 157684)
Lamerdin, J.E., McCreedy, P.M., Skowronski, E., Adamson, A.W.,
Burkhardt-Schultz, K., Gordon, L., Kyle, A., Ramirez, M., Stillwagen, S.,
Garner, J., Dangnan, L., Poundstone, P., Christensen, M.,
Georgescu, A., Ayala, J., Liu, S., Bruce, R., Quan, G., Montgomery, M.,
Ow, D., Nolan, M., Trong, S., Kobayashi, A., Olsen, A.O. and
Carriano, A.V.
Unpublished
JOURNAL 2 (bases 1 to 157684)
REFERENCE Lamerdin, J.E.
AUTHORS Direct Submission
JOURNAL Submitted (31-JAN-1998) Joint Genome Institute, Lawrence Livermore
National Laboratory, 7000 East Ave., Livermore, CA 94551, USA
FEATURES
source location/Qualifiers
1. 157684
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="BAC-10817"
/cell_line="CJ7"
/lab_host="HS996 (modified DH10B/r)"
/notes="BAC clone obtained from Genome Systems. Library
made from cell line CJ7, derived from mouse strain 129SV."
/rpt_family="L1"
649..802
repeat_region
/rpt_family="L1"
676..753
misc_feature
/rpt_family="L1PAB"
/notes="predicted exon, program: grail2exons_human_1.3,
frame: 0, quality: excellent, score: 86.000"
1516..1563
repeat_region
/rpt_family="(GGA)n"
complement(1380..3208)
repeat_region
/rpt_family="POLY_A"
complement(4735..4991)
misc_feature
/notes="predicted exon, program: grail2exons_human_1.3,
frame: 1, quality: good, score: 52.000"
complement(5770..6199)
misc_feature
/notes="BLASTN similarity to U17089 (72..503); match: 0.93,
score: 8.5e-139; database searched: nt; Mus musculus MT
transposon-like element, clone MT17."
complement(6121..6200)
repeat_region
/rpt_family="THER1-INTERNA"
9771..9823
repeat_region
/rpt_family="(CA)n"
complement(9823..9908)
misc_feature
/notes="predicted exon, program: grail2exons_human_1.3,
frame: 2, quality: excellent, score: 86.000"
10150..10173
repeat_region
/rpt_family="(POLY_A)"
complement(10838..10877)
repeat_region
/rpt_family="AT-rich"
complement(11272..11610)
misc_feature
/notes="predicted exon, program: grail2exons_human_1.3,
frame: 1, quality: good, score: 51.000"
complement(12688..12820)
repeat_region
/rpt_family="FRAM_C"
complement(13313..13358)
repeat_region
/rpt_family="AT-rich"
complement(13865..13969)
repeat_region

```

```

repeat_region
/rpt_family="L1MC4"
complement(14440..14884)
misc_feature
/rpt_family="L1MD3"
14681..14844
/notes="predicted exon, program: grail2exons_human_1.3,
frame: 1, quality: good, score: 67.000"
complement(15549..15632)
repeat_region
/rpt_family="AluUo"
complement(16026..16058)
repeat_region
/rpt_family="(CAA)n"
complement(16233..16292)
repeat_region
/rpt_family="(CA)n"
16282..16419
/rpt_family="(GA)n"
complement(17364..17432)
repeat_region
/rpt_family="AT-rich"
18848..18901
/rpt_family="(CA)n"
18906..19029
/rpt_family="(TAA)n"
19108..19156
repeat_region
/rpt_family="AluUb"
complement(19158..19272)
repeat_region
/rpt_family="(TAA)n"
19627..19758
/rpt_family="(TGA)n"
complement(20957..21319)
/notes="Ribosomal pseudogene, --DPS similarity to
g1133881|sp|P12947|RL31_HUMAN 60S RIBOSOMAL PROTEIN L31;
g1171328|p|t|RS3H031 ribosomal protein L31 - human;
g1171329|p|t|RSRT31 ribosomal protein L31 - rat; g136130
(X15940) ribosomal protein L31 (AA 1-125) [Homo
sapiens]; g1157115 (X04809) ribosomal protein L31 (AA
1-125) [Rattus norvegicus]; g1163285|gnl|P1D1e276436
(L281183) ribosomal protein L31 [Sus scrofa]. Score: 405
identity: 91/125 (72%)."
22316..22474
/rpt_family="AluUb"
23291..23335
repeat_region
/rpt_family="Alu"
23316..23350
/notes="Unresolved number of A's in homopolymeric A-
region."
23336..23434
/rpt_family="(POLY_A)"
23728..23867
repeat_region
/rpt_family="AluUo"
23874..23925
/rpt_family="(GA)n"
24688..24798
misc_feature
/notes="predicted exon, program: grail2exons_human_1.3,
frame: 2, quality: good, score: 52.000"
24883..25000
repeat_region
/rpt_family="(GA)n"
complement(25229..25300)
repeat_region
/rpt_family="7SLRNA"
complement(25355..25449)
repeat_region
/rpt_family="7SLRNA"
complement(25350..25465)
repeat_region
/rpt_family="AluSp"
25477..25637
/rpt_family="AluSg/X"
27395..27440
/rpt_family="(CA)n"
27440..27585
repeat_region
/notes="low confidence in sequence within low complexity
repeat region."
27441..27566
repeat_region
/rpt_family="(CA)n"
27598..27742
repeat_region
/rpt_family="(CA)n"
complement(28223..28355)
/rpt_family="AluSp"

```

OY	4230		gctgcgaataagatgtggcaccacacacccagttagttagtttccttgcttact	4289
Db	154763	GCTGGGATTAAAGCGGTGTGCACCATCCGCCGGTCCTAATTTTATAATGTAACATTGG	154704	
OY	4290	tctcat	4295	
		- -		
Db	154703	TCTTTT	154698	
RESULT	15			
LOCUS	AC004155			
DEFINITION	AC004155 128026 bp DNA ROD		28-MAR-1998	
	Mus musculus DNA from BAC 10818 containing the Ercc-4 gene,			
	complete sequence.			
ACCESSION	AC004155			
NID	92896800			
VERSION	AC004155.1	GI:2896800		
KEYWORDS	HTG.			
SOURCE	house mouse.			
ORGANISM	Mus musculus			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;			
AUTHORS	Buthenia; Kodenita; Sciurognath; Muridae; Murinae; Mus.			
	Lamerdin,J.E., McCreedy,P.M., Skowronski,E., Adamson,A.W.,			
	1 (bases 1 to 128026)			
	Burkhardt-Schultz,K., Gordon,L., Kyle,A., Ramirez,M., Stillwagen,S.,			
	Pian,H., Velasco,N., Ganes,J., Dangnanan,L., Poundstone,P.,			
	Christensen,M., Georgescu,A., Ayala,J., Liu,S., Attix,C.,			
	Andrzejewski,T., Frankheim,M., Antico-Keller,G., Coeffield,J., Duarte,S.,			
	Lucas,S., Bruce,R., Thomas,P., Quan,G., Kronmiller,B., Arellano,A.,			
	Montgomery,M., Ow,D., Nolan,M., Trong,S., Kobayashi,A., Olsen,A.O.			
	and Carraro,A.V.			
TITLE	Sequence analysis of a mouse EAC containing the DNA repair gene			
JOURNAL	Ercc-4 (xpr)			
AUTHORS	Unpublished			
JOURNAL	2 (bases 1 to 128026)			
TITLE	Lamerdin,J.E.			
JOURNAL	Submitted (19-FEB-1998) Joint Genome Institute, Lawrence Livermore			
FEATURES	National Laboratory, 7000 East Ave., Livermore, CA 94551, USA			
Source	Location/Qualifiers			
	1..128026			
	/organism="Mus musculus"			
	/db_xref="taxon:10090"			
	/clone="10818"			
	/cell_line="CJ7"			
	/lab_host="HS996 (modified DH10B/r)"			
	/note="Clone 10818 containing mouse Ercc-4 obtained from			
	Genome Systems. CJ7 cell line is derived from mouse strain			
	129SV."			
	23..52			
	/rpt_family="(CAAAA)n"			
	complement(489..532)			
	/rpt_family="(CCAA)n"			
	complement(544..596)			
	/rpt_family="(Alus"			
	629..661			
	/note="predicted exon, program: gtrial2exons_mouse_1.3,			
	frame: 1, quality: gcod, score: 66.000"			
	complement(712..742)			
	/rpt_family="(AT_rich"			
	complement(1415..1582)			
	/rpt_family="(CA)n"			
	3254..3316			
	/rpt_family="(Alu/FAM"			
	3516..3576			
	/rpt_family="(CA)n"			
	complement(3882..3912)			
	/rpt_family="(CAA)n"			
	4480..4601			
	/rpt_family="(AAAA)n"			
	4706..4824			
	/rpt_family="(GGGA)n"			


```

23663. 23807
/repeat_region /rpt_family="(GA)n"
23842. 24064
/repeat_region /rpt_family="(GGA)n"
25418. 25485
/repeat_region /rpt_family="(MTTTF"
complement(25789. 25815)
/rpt_family="(AT_rich"
complement(25863. 25937)
/rpt_family="(AluSp"
complement(27004. 27134)
/repeat_region /rpt_family="(TA)n"
27997. 28083
/rpt_family="(TAA)n"
28352. 28433
/rpt_family="(TAA)n"
28762. 28910
/rpt_family="(CAT)n"
/note="Predicted exon, program: grallexons_mouse_1.3,
frame: 1, quality: excellent, score: 76.000"
29273. 29397
/note="Predicted exon, program: grallexons_mouse_1.3,
frame: 0, quality: excellent, score: 92.000"
29734. 29758
/rpt_family="(CA)n"
complement(29885. 30026)
/rpt_family="(AluJo/ERAM"
complement(30337. 30398)
/rpt_family="(TRGA)n"
complement(32477. 32523)
/rpt_family="(CA)n"
32625. 32681
/rpt_family="(RNA-Cys-TGC"
32766. 32828
/rpt_family="(POLY A"
33036. 34140
/rpt_family="(LMA2"
complement(34829. 35395)
/rpt_family="(HERVR"
complement(35279. 35633)
/note="BLASTX similarity to P21414 (817. 911); match: 0.
score: 2.3e-178; database searched: nr; POL POLYPROTEIN
(Protase, REVERSE TRANSCRIPTASE"
complement(35714. 35770)
/note="BLASTX similarity to P21414 (709. 727); match:
0.73, score: 2.3e-178; database searched: nr; POL
POLYPROTEIN (Protase, REVERSE TRANSCRIPTASE"
complement(35765. 35979)
/rpt_family="(HERVC2"
complement(35985. 36156)
/note="BLASTX similarity to P21414 (545. 600); match:
0.73, score: 2.3e-178; database searched: nr; POL
POLYPROTEIN (Protase, REVERSE TRANSCRIPTASE"
complement(36157. 36532)
/rpt_family="(HERV"
complement(36549. 36695)
/note="BLASTX similarity to P21414 (354. 402); match:
0.55, score: 2.3e-178; database searched: nr; POL
POLYPROTEIN (Protase, REVERSE TRANSCRIPTASE"
complement(36728. 36837)
/rpt_family="(GA)n"
complement(37079. 37151)
/note="Predicted exon, program: grallexons_mouse_1.3,
frame: 1, quality: excellent, score: 100.000"
/note="Predicted exon, program: grallexons_mouse_1.3,
frame: 2, quality: good, score: 65.000"

```

Tue Sep 21 09:50:17 1999

us-09-037-657-38.rge

Page 22

[illegible]

Search completed: September 19, 1999, 16:18:57
Job time: 16851 sec

OM nucleic - nucleic search, using sw model

Run on: September 19, 1999, 11:40:06 ; Search time 244.65 seconds

(Without alignments)
12100.032 Million cell updates/sec

Title: US-09-037-657-38
Portfolio: 11033

Sequence: 1 gcgagccgtgcagtgtatc.....tatgtgctggggtggggga 11832

Scoring table: IDENTITY_NUC

Searched: 311585 seqs, 125096042 residues

Database : N_Geneseq_36:★

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query	Match	Length	DB	ID	Description
C	1	11830.4	100.0	11832	1	V27148	Nucleotide sequenc	
C	2	6561.6	55.5	6663	1	V27145	Nucleotide sequenc	
C	3	737.2	6.2	1930	1	V27158	Unspliced murine N	
C	4	737.2	6.2	1938	1	V27142	Novel haemopoietin	
C	5	280.4	2.4	1639	1	V27140	Novel haemopoietin	
C	6	280.4	2.4	1673	1	V27141	Novel haemopoietin	
C	7	277.2	2.3	1656	1	V41688	Nucleotide sequenc	
C	8	254.4	2.2	1634	1	V27143	Nucleotide sequenc	
C	9	248	2.1	1724	1	V70896	CDNA encoding rat	
C	10	221.8	1.9	1690	1	V70894	CDNA encoding huma	
C	11	221.8	1.9	1813	1	V70895	CDNA encoding a	
C	12	220.2	1.9	1879	1	V41689	Nucleotide sequenc	
C	13	190.8	1.6	1391	1	V27144	Nucleotide sequenc	
C	14	141.2	1.2	560	1	V27159	PCR product for hu	
C	15	121.2	1.0	143	1	V27157	Nucleotide sequenc	
C	16	117.6	1.0	2509	1	V07515	Vascular endothell	
C	17	111.6	0.9	4688	1	O32370	p1A genomic DNA, N	
C	18	111.6	0.9	4688	1	O72475	Enfite p1A genomid	
C	19	111.6	0.9	4658	1	T01176	B15A antigen prec	
C	20	111.6	0.9	4658	1	T14692	MAGE genomic DNA.	
C	21	108.4	0.9	4105	1	X23316	Mouse 1-alpha-OHase	
C	22	108.4	0.9	1688	1	X23318	Mouse 1-alpha-OHase	
C	23	107.4	0.9	2554	1	N90724	Nucleotide sequenc	
C	24	107	0.9	229	1	V70897	Expressed sequence	
C	25	106.6	0.9	6727	1	T88014	Expressed sequence	
C	26	106.2	0.9	3360	1	V19608	Murine IL-5 CDNA g	
C	27	105.4	0.9	14180	1	V05159	Mus musculus EPCR	
C	28	103.6	0.9	3283	1	O13855	Mouse butyrophilin	
C	29	103.4	0.9	4923	1	T34501	Murine GCSF recept	
C	30	100.8	0.9	210	1	V70899	RHAMM 1 gene intro	
C	31	100.6	0.9	1311	1	V33155	Expressed sequence	
C	32	100.2	0.8	6259	1	T86366	Mouse neutrophil e	
C	33	100.2	0.8	3892	1	V34385	SM22 protein gene	
C	34	100	0.8	2190	1	T86367	Upstream region of	
C	35	100	0.8	1681	1	T86368	SM22 protein gene	
C	36	100	0.8	785	1	T86369	SM22 protein gene	
C	37	100	0.8	2147	1	V22706	SM22 protein gene	
C	38	99.8	0.8	535	1	V09117	Mouse recombinase	
C	39	99.6	0.8	3364	1	V07266	5' nucleotide sequ	
C	40	98.4	0.8	4989	1	O85962	Mouse creatine kin	
C	41	97	0.8	5000	1	V33688	Mouse STRL gene. A	
C	42	96.6	0.8	5000	1	V33755	Mus musculus SOCS1	
C	43	96.2	0.8	3419	1	X28077	Mouse neutrophil e	
C	4	96.2	0.8	3419	1	X28077	HS25T coding sequen	

C	44	95.8	0.8	1355	1	Q22971	Sequence encoding
c	45	95.2	0.8	277	1	X27006	Differentially exp

ALIGNMENTS

```

RESULT      1
V27148
ID V27148 standard; DNA; 11832 BP.
AC V27148;
DT 29-SEP-1998 (first entry)
DE Nucleotide sequence for murine NR6 containing additional 5' sequence.
KW Haemopoietin receptor; cell proliferation; cell differentiation; cancer;
   cell survival; therapeutic; neuronal proliferation; drug screening; ss;
KM Mouse.
OS Mus sp.
PN MO981125-A2.
PD 19-MAR-1998.
PF 11-SEP-1997; G02479.
PR 11-SEP-1996; AD-002246.
PA (AMRA-) AMRAD OPERATIONS PTY LTD.
   (DZIE/) DZIELEWSKA H E.
PI Alexander W. Fabry L, Farley A, Hilton DJ, Kikuchi Y,
   PI Kojima T, Meda M, Nash A, Nicola NA, Nakar S, Willson T,
   PS Zhang J;
   DR WPI; 98-260970/23.
PT New isolated haemopoietin receptor - used for developing products
   PT for modulating proliferation, differentiation and survival of cells,
   PT e.g. neuronal cells
PS Claim 9; Fig 3; 182pp; English.
CC The NR6 protein is a novel Haemopoietin receptor (HR). Interaction
CC between the novel HR and a ligand facilitates proliferation,
CC differentiation and survival of a wide variety of cells. The HR and its
CC derivatives can be used for modulating the activity of the receptors e.g.
CC to regulate development, maintenance or regeneration in an array of
CC different cells and tissues in vitro and in vivo. They can be present in
CC therapeutics used for modulating neuronal proliferation, differentiation
CC and survival. The products can also be used for detection and diagnosis,
CC e.g. for cancers or predisposition to cancers, or for drug screening.
SQ Sequence 11832 BP; 2447 A; 3367 C; 3298 G; 2720 T;

```


Query Match	100.0%; Score 11830.4;	DB 1;	Length 11832;
Best Local Similarity	100.0%;	Pred. No. 0;	
Matches 11831;	Conservative 0;	Mismatches 1;	Indels 0;
Gaps			0
OY 1 gcgcgctgcagtgattactcaacgcgttgcgaccaccaaccgcggcgctagtatga 60			
Db 1 GCGCCTGCGAGTGATTACTCAACGGGTGGCGACCCACCACGCCGGCGTGAATGGA 60			
OY 61 ttttcctgttggggaaagttaagaagtatttgggaacctttctgcacgatggaaacta 120			
Db 61 TTTTTCGCTGGGGGAAAGTAAGAAGTTTAGGGGAACCTTTCTGCACCATGGGAACATA 120			
OY 121 gaaatcaggagtttgtcccttcctcccacaagacaacccctccccataagccaactcat 180			
Db 121 GGATCGAAGGTTTCGCTCCGCTTCCCNAAGACACACCTCTCCCCATTAAGCCCACTCAT 180			
OY 181 aaaggtccctctgcacgcgtctcggagacatcccatacacatatcccgacagatatgatagt 240			
Db 181 AAGGCTCCTCTGCACGGGCTCCGGGACATCCCATATCAATACCCGACAGATATGATAGT 240			
OY 241 tgagaaaggagaccagagggccggagagactccctccctgccttttgatcttccccccctctgc 300			
Db 241 TGAAAAGGGACACAAGGCCGAGACTCTCCCTTCCTTGTGCTTTCCCTCCCTCCCTCTGC 300			
OY 301 acgaaagaaactaacacagcagatggagagagtgatgaagccttagaggttgagatcgatag 360			
Db 301 ACGAAAGGAAGACTACACCGCATGGAGAGAGTGGCATGAAGCCTTAGAGGTGGGATCGGATAG 360			
OY 361 gaaccatgcaccccaagagaaaggagacttgatgcacatttcacaaactctctgggaaggagaaga 420			

Db	361	GACCCATGCAACCCAGAAAGGAGACTGTGTGTCGACACTTTAAACTCTCTTGCGAAGAA	420
QY	421	aggcgctgaagaaatgaacggggtccaggtactgtctcaatgtgtgtgtgcygaaacaa	480
Db	421	AGGCGCTGAAGAGATGATCAAGCGGCTCAGGTACTGCTCATATGTGTGTGCGGACCAAG	480
QY	481	tgggtatcgggggccccctgaagaggggcgggaaagtgtgatatgaaagatcccgtagact	540
Db	481	TGGGTATGCGGGGCCCTTAAGAGGGGGGGGAAAGTGTGATAGAAAGATCCCGTAACT	540
QY	541	ggaaggatccctgtgaaagacacaggggtctcggagctaggaaacccattcgaattaaagt	600
Db	541	GGAGGGGATCTTGGAAAAACACAGGGGCTCGAGCTAGGAAACCATTCCGATTTAAGGT	600
QY	601	acaagatccacagatgataggggggtgtggaagccctgtgaaagggcgagaccaaagaaaggtcc	660
Db	601	ACAGATATCCAGATGAGAGGGGTGGAAAGCTTGGAGAGGGGCGGACAGAGAGAGTCC	660
QY	661	caaggctgtgtgtgggaaagagtgtgggggtctcgcgcaagagatvggaactcagaagtgtg	720
Db	661	CACGGCGTGTGTGGAAAGTAGTGGGGGCTTCCGCGAGAGAGATGGAGATTCAAGATG	720
QY	721	gtaaacttggcggagcgcgcgccggggcggggtgtgcgcgcgcgcgcgcgcgcgcgcgcgc	780
Db	721	GTAATCTGGCGGAGCCCG	780
QY	781	gtcgggggcccaacatcaaccccccccaagggcgcgcgccggcgccggggcgcgggcgcg	840
Db	781	GTCGGGGCCCAAGATTAACCCCCCCCAAGGGCCCGGCGCGGGCGGGGCGGGGCGGGG	840
QY	841	cggggcgaagcggcgcalttagcgcctgtgtcaatttcgctcgtcacaactgtgtccgcgctt	900
Db	841	CGGGGCGAGCGGCATTATACGCGCTGTCTAATTGCGGTGCTACAGTGTGCTCGGCGCTT	900
QY	901	cgctgtccgcgcgcgaatgaaagcgcggtgaaagaccggaacccaattgtcaaccccgagact	960
Db	901	CGCTGTCCGGCCCAAGTGAAGCGCGGTGAAGACCCGAGCCCATCTGCAACCCCGAGACT	960
QY	961	cgccccccccaataacaggcggtgtgaagtaacggccccgtgttgccgcgcaccccccaatgc	1020
Db	961	CGCCCCCGCCCAATACGGGGGTGCAATCACGGCCGTTGCGGCGCACCCCCATATCCGCG	1020
QY	1021	gggtgcgcgggccccgtctgcacccaatccgacgacggcgccggcgcgcgctgtctccgct	1080
Db	1021	GGGTGCGCCGGGCCCCGTCGCCCAATCCGGGGGGCGCGCGCGCGCGCTGTCTCCGCT	1080
QY	1081	gtgtgtgcctctgtctctgtctgtctgtccctcggggtgtgcctggggcggaatccggatga	1140
Db	1081	GTTGTGCTCTGTGTCTGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT	1140
QY	1141	gtaccggtggccccgtcccccaactcccccaaggaaagcgggttccggcgcccccggggggt	1200
Db	1141	GTAACGTGCGCCTGTGCTCCCCACCTCCCAAGGAAACCGGGAATCCGGCGCCCCGGGGGT	1200
QY	1201	agtcgcggggatgtgaagaaagggggcgcgacacctgtgaagtcctccggaaacaaagaa	1260
Db	1201	AGTCGCGGGGATGGAAGAAAGGGGCGCGACCGCACTGAGAGTCCCGGAACAAAGAA	1260
QY	1261	ggcggccttcggggcgccctcacctgttggggtcatalgtgcacacacacccagctcccaag	1320
Db	1261	GGCGGCCCTTGGGGCGCCCTCACCTGTGGGGGTCAATGACCAACCAAGCTCCCAAG	1320
QY	1321	agttccccgttataacatcagaaggcccttatctgtatcccccttgcgagagctgtcttgcc	1380
Db	1321	AGTATCCCGTTATATACATCAAGGCGCTTATCTGTATCTCCCTTTCGAGGCTGTGGCC	1380
QY	1381	aggtcagtttgaagagacatcgacagtgctctggaaacccccctctcaaggtgtgtcgagac	1440
Db	1381	AGGCTACGTTGAAGAGACATCGAGTGTCTGGAGACCCCGCTCTTCAAGGTGTGGAC	1440
QY	1441	gcttcggggcgacgcgtgtgtctgtgatatcaagcggaaagaaagcctccctgtgcgg	1500
Db	1441	GCTTCGGGGCGACGCTGTGTCTGTGATATAGACCGGAAGGAAGCTCCCTGTGGCGG	1500

[illegible]

[illegible]

Db	3661	CTGGGAAGAACTCAATTCAGGGGCCAGAGCTGAACCAATGTGGTGCCACCTCATCATCTG	3720
OY	3721	gggaatgaaggaacgctcagaagcctgaagtttaactcttaataatgaagaaactggggggtg	3780
Db	3721	GGGCAATGAAGAACCGTCTCAGGGCTGCAGTTTACCTCTTAATGGAACCTGGGGGGGT	3780
OY	3781	gaagcctctgtcttcgagcctctcttgaaatcgattttgtttttgttttttc	3840
Db	3781	GCAGCCTGTGTTCTCCAGGCGCTTTTGGAATGGTTTGTTTTGTTTTGTTTTTC	3840
OY	3841	ataactcttctcctcaatcccaatccgggaactgtttccctcaagggttgaagacc	3900
Db	3841	AATATCTTTTTCCTCATATCCATCCACCGGAGACTGTTCCTCCCTAAGGGTTGAGAGCC	3900
OY	3901	tgcagctctcccaacttttcttgtttttatcccaaggagcctttgcaatgaatgccca	3960
Db	3901	TGCACTCTTCCCTAACCTTTTCTTGCTTTTACCCAGGGCCCTTTCACATGAAGTCCA	3960
OY	3961	cctctccctctggccaactgagggtccagccttaactgcatitgtcctcttgtaactgtcc	4020
Db	3961	CCTGCCCCCTTGGCCAACTGGGGCTCAGGCTTACAGCATTTGGCTTTGGTAAGTGC	4020
OY	4021	caggcctctcttgacacagaggtttagcccaagctccctctctctctcccccctt	4080
Db	4021	CAGGGCCTCTCTGCACACAGAGGTGAGCCAGCTCCGCTCTCTCTCCGCCCTT	4080
OY	4081	ctctttgtctctgaagcttaatttttttttttttttttttttttttttttttttttttt	4140
Db	4081	CTCTTTGTCTCTGAGCTTAATTTTTTTTTTTTTTTTTTTTTTTTGGCTTTTGAACAGGGT	4140
OY	4141	tctctgtacagccctctgctcgcctctggaactcattctgtacagagctagactcaaac	4200
Db	4141	TCTGTGTAAGCCCTTGCGTCCCTGGACATCTGTGTAACAGGCTAGCCTCAACT	4200
OY	4201	acaaactactcgtcctctctgctttccagttgtgcactaaagaatgtgggccaaccaact	4260
Db	4201	ACAAACCACTGCTGCTGCTGCTTCCAGTCTGGACACTTAAGATGGGGCCACACAACT	4260
OY	4261	agtagttaaggttttctgtgttcttatctcctatagtaactcagctcagctctgtgcatatg	4320
Db	4261	AGTAGTTAAGTGTTCCTGTGCTGCTTATTCCTAATGATGTAACCTCAAGTTCTTGCAATATG	4320
OY	4321	tagcgatgagatgagataaagagatgagatgagatgagatgagatgagatgagatgagatg	4380
Db	4321	TAGCGATGAGATGAGATAAAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG	4380
OY	4381	cttgaatcgtctcgtagtgaaanaaagaaactcagagaaactgaatgagatgagttccag	4440
Db	4381	CTTGAATCGTCTGATGAAANAAGAAAGAACTCAGAACTGAATGAGATTAGTTCACAG	4440
OY	4441	ggagcctctgctcgtcgtgttctcatgtggagctccctgtgaacttcccacaactccac	4500
Db	4441	GGCAGCCTGGCCCTGCTGCTGCTCATGAGGAGCTCCCTGTGAATTTCCCCACACCTCCAC	4500
OY	4501	caacctgcactcgtgtgtgtctgaagaagaagccaatggccaagtggggaacaagactc	4560
Db	4501	CACCTGCAATCTGTGTGCTACAGAAAGGCCAATGGCCAGATGGGGACACAGACTC	4560
OY	4561	agggaagcttggaatatgttcccccctcctacatacctagaagccttgttgcctcctgaagc	4620
Db	4561	AGGGAAGCTTGGAAATATGTTCCCTCCTCATATCCAGGCTGTGTGCCCTGAGAGGC	4620
OY	4621	ccagcctatagatgaagggcagcgtgtggctgtccctaagtgttgtagcagaagaagggtg	4680
Db	4621	CCAGCCTATAGTAGGAGGCAAGCTGTGGGCTCTAAGATTGGGTAGCCAAAGAGGGGTG	4680
OY	4681	gtccctaaagggtgtgtccagaggttgaagttcctcaaatgtggcctacaagtgccct	4740
Db	4681	GTCCTCAGGGGTGGGTACAGAGATTGAGTCAATTCCAAAGTGGCCATCAAGTGGCCCT	4740
OY	4741	aggaatgattgtggaagatcagaactcctctgttggaatttgaagggccttgatgtg	4800

D	b	4741	AGGAAATGATTTGGAGAGTCCAGAACTCCTGTTGGAGTTGTAGAGGGCCCTGCATGTGG	4800
Q	y	4801	gctctctggtcgtgcctctctctctctgctcttgacagtcacctcggtgtgtgtgga	4860
D	b	4801	GCTTCTGTGGCGTGCCTCTCTCTGTGTGTCCTTTGCACAGTCCCTCGTGTGTCTGGGA	4860
Q	y	4861	tgttagaaggacacggyggaataatgaagctcaagccctcagctgcctcctcaagttac	4920
D	b	4861	TGTAGAGAGGGCACGGGGAAATGAAGGCTCAGACCCCTCAGCTTCCCTTCAGGTTAC	4920
Q	y	4921	ccaaagaaggctaacctctctctctgagcaggtctcactgtatgacagaattggtctaca	4980
D	b	4921	CCAAAGAGGGCTCACCTCTCTCTGGACAGGCTCTCACTGTATGACAGATTGGCCCTACA	4980
Q	y	4981	tttgatccctctctctgctcctctgagtagacaacattcacagggtgaagatttaca	5040
D	b	4981	TTTGATTCCTCTCTCTTGGTCTCTCTGGATGCAAAACATTTCACAGGTAGGATTTTACA	5040
Q	y	5041	ttttaatatgtccatctctccagaacaacactgttaggtttaaggtatagtgaagagac	5100
D	b	5041	TTTTAGATATGTCCATCTCTCCAGAAACACACTGTAGGTGTGGGTATCAGTAAGTAAGAC	5100
Q	y	5101	accccccagcagacaaagaattggagggaggaagaattgtgtaaccagggccatgtctga	5160
D	b	5101	ACCCACCAGGACAGCAAAAGATTGGAGAGGAAGAATTGTATTAACCCAGGCCATCTCTTGA	5160
Q	y	5161	tgtgcttatgtatacccgaactctggagcgtctgaagcagagagattccaaagtccaaga	5220
D	b	5161	TGGCTTATGTGTAACTCCCAAACTCTGGACGCTGAGGCGAGGAGATTCCAAAGTTTCAAGA	5220
Q	y	5221	cagtggtctcaggtaaatgagaccctgtcaagaagaagaataaagaagacaagaata	5280
D	b	5221	CAGTGTGTCTAGGTATGAGACCCGTCTCAAGAAAGAAATAAATAAGACACAGAA	5280
Q	y	5281	atgtttataagcctgtgtagaagactgtgtgtgtaagggcactctgcctcaataagaatga	5340
D	b	5281	ATGTTTATAGCGCTGTGAGACAGCTTGTGTGGTGAAGGGCGACTTGCCTCCATCAAGAGGA	5340
Q	y	5341	cctcagcccccactccctcagatgcacatcgtgtgtagaaggagagaagaactccagctgtgac	5400
D	b	5341	CCTAGGCCCATCCTCTAGGAATTCATATGTATGAAGAGAAACAAACTCCAGCTGCTTGCAC	5400
Q	y	5401	tccatacatgtctcccaatgtcacacacacagggagacataatcaatataagatga	5460
D	b	5401	TCCATACATGTGCGCCAAATGTGCACACACACAGGGAGACATTAATTAATAGAGTATA	5460
Q	y	5461	tttgcttagatttgatgtaggaacattatgacatgattttcaaatcttattatgatttat	5520
D	b	5461	TTTCTTATAGATTGTAGTAGAGCATTTATGATCATGTATTTAAATTTTATTTGATTTTAT	5520
Q	y	5521	gaaatataccctttgtattgtttgtttgtttgtttgtttgtttgtttgtttgtttgtttgt	5580
D	b	5521	GAATAATTAACCTGTTTGTATTTATTTGTTTGTGTTTGTGTTTGTGTTTGTATTTTGAACAGG	5580
Q	y	5581	gctctctgtctagtcctgtcgtctgctcttgaaactcaactctgtgagacsaagctgtgcctga	5640
D	b	5581	GCTTCTGTGTAGTGCCTGCGCTGCTCTTGTGAATCACTGTATGACACAGCGCTGCGCTTGA	5640
Q	y	5641	actagaagaatcgcgcgtctgtgtctcccaagtgcttagattaaaggtgtgcactgcac	5700
D	b	5641	ACTAGAGAAATCCGCGCTGCTGTGCTTCCCAAGGCTTAAATTAAGGTGTGCACGTGCAT	5700
Q	y	5701	tcagcaaaattgtatactttaaccccaagatttgggaagcagaagcagaactaaatgttga	5760
D	b	5701	TCACAAAATTTGCTATCTTTAACCCCAAGATTTTGGAGGCAAGGCAACATTAATGTGGA	5760
Q	y	5761	attccagggctagccaagatacagagttagaaccttatcttcaaccctcccccacaacac	5820
D	b	5761	ATTCAGAGCTAGCCAAAGATACAGATGTAGAGCCATTTTACCTCCCTCCCAAAAC	5820
Q	y	5821	ccaaatgtatatttgctctgtgtatgtatcatatgtgtgtgtgcagacgtaaatgtccaag	5880
D	b	5821	CCAAATGTATTTTGTCTCTGTGTATGTATCATGTGTGTGTGCACGTAATGTCCAAG	5880

QY	5881	acaactgtgaaagtctctctcgltccaaagctctaaagctcgaatccaactaaaggtctc	5940
Db	5881	ACAACCTGTGAAAGTCTCTCCCTCCCTACAGCTCAAGTCTGATTCATTAACAACTAAGTCTCTC	5940
QY	5941	aggcttagcccaactctctcttcttctgacagagcaattcaactcgtgcccgttgtaactgat	6000
Db	5941	AGGCTTAGCCCAACTCTCTTCTTATGTACTGAGCCATTTCAGTGGCCCTGGATTGACTGAT	6000
QY	6001	gaattaaattttgagataaaggtctctctgttagctctagctaggctcaactatgaactccc	6060
Db	6001	GAATTAATTTTGTGATTAAGGCTCTTGTAGCTCTAGTAGAGCTCAAACTATGAACCTCC	6060
QY	6061	aaggtacatttgtagctgtgtgtactcttgcttccaaacccaagtggtgaaatgataactag	6120
Db	6061	AAGGTACATCTGAGCTCTGTACTGTCTTCTTCACCCCAACTGTGTGAATGATACTAG	6120
QY	6121	gcacacactctctctgaggaaagggtctgtccctgtgacctgaatttctgtgctccagctcaat	6180
Db	6121	GCACACACTCTCTGAGGAAAGGGCTGTGCCCTGTATTTGTGTGCTCAGCTTCAT	6180
QY	6181	gagtgctttgggtctcgltgttcttcttcttatctgttgaatrggttgaacacctgttaa	6240
Db	6181	GAGTGCTTGGGTCTCGTGTGTTCTTTCTTTACTGTATCTGTGAAGGGGTGAACACCTGTTCAA	6240
QY	6241	gacttcttcgactctctgaaatactcagagaggttgaaggagcttgaagtggtctcaatcccat	6300
Db	6241	GACTTCTCGACTCTTGAATACTCCAGGAGGGTGAAGGACTTGAAGTGGCTCATCCAT	6300
QY	6301	gccttaacaaagtgtcgtctcttctgaccccaagacaaactgttaatacagcccccaagacccac	6360
Db	6301	GCCTAACAAGTGTGCTCTTGTGACCCCAAGACAAAGCTGTAATCAGCCCCCAGAACCCAC	6360
QY	6361	ccctctatcggtctccctccctctgaagctaacctcgtctctataacttgaaagacaaacttgggc	6420
Db	6361	CCCTCTATCGGCTCCCTCCCTGGAACCTCAGCTCTCTATACATGGAAGAACACCTGGGGC	6420
QY	6421	caacgctgaggggtctcactacagaccttcaatgtgcgcgcgcgcctctgagctgtccgc	6480
Db	6421	CACGCGTAGGGGGCTGTACTAGTGAACCTTAATGATGGCCGCCCTGTAGCTTCCG	6480
QY	6481	ccctccctaaacctcacaacctgtgcctgtgcctgtgtaaccttaatlvgttccagagca	6540
Db	6481	CCCTCCCTAAACCTCCACCCCTGTGCCCTGTGCCCTGCTTAACCTTAATGGGTCTCAGGACCA	6540
QY	6541	gtcagagagacaaatctgtgtgtgtacagcccgagaaacagacatctcgtgtgtcctgcct	6600
Db	6541	GTCAGAGACAAATGTGTGTGTGTACCCCGAGAGGGCACCAATCTTGCGTGGCTCTGCT	6600
QY	6601	ctatgtctgtcgttaagtgtggggcccaagacactcagagatagatagatgtgggglttgcaatgtaca	6660
Db	6601	CTAGTGTGGCTGTAAAGTGGGGCCCCAGACACTCAGAGATAGATGGGGGTGGCAATGACA	6660
QY	6661	gatttaagcctgtgtgtcttctgtccctggggcagagacatgggtctctcaactgtcatgacag	6720
Db	6661	GATTTAAGCCTGTGTTCTTGTGCTGGGGAGAGGCAAGGGGCTCTCACTTGCAATGACG	6720
QY	6721	catgtcatatcccaagcacagcaatgcaactctagaggaagctgtgtgtacacgtgtcccc	6780
Db	6721	CATGTCTATACCCAGACAGCAAGCAATTGCACACTAGAGGAAGCTGTGTGCTGCACTGTCCC	6780
QY	6781	tgtgtaccccaagcctttagaagaagctgtcatatgttctctgttagtgtccccctgagaagc	6840
Db	6781	TGTGTACCCACAGCCTTAAGAAAGCTGTCAATGATTTCTTCTTATGTGCCCTCAGAGAC	6840
QY	6841	ccttaacatacaagctgtgtgtcccggaatagaagaatctacgtgtgcgcgtggagacag	6900
Db	6841	CCTTAACATCAAGCTGTGTGTCCCGGAATAAGAGATCTACGTGCGCGCTGGACACCG	6900
QY	6901	gtgcacacggggagaaattctacataccaactactccctcaagtataagcttgagttgtg	6960
Db	6901	GTGCACACGGGGAGAAATTTCTACATACCAACTACTCCCTCAAGTACAGCTGAGGTGG	6960

QY	6961	taccacgccaagccctctgtgacgtctcggaactccctctctcgatcaaatg	7020
Db	6961	TACCACGCAAGCCCTCTGTGACTCTCGGAACTACTTCCTTCTGATCAAAATG	7020
QY	7021	tctcgtttatgaaactcaaaaggagactctgcaccctccacaggtgtagtcagaataa	7080
Db	7021	TTCTCGTTTATGAACCTCAAAAGGAGCTCTGCACCTCCACAGAGTGATGGTACGATGA	7080
QY	7081	caactgtggaagagatcaccaactctgaggccctcaactatgcatatcccaaggccggc	7140
Db	7081	CACATGTGAGAGATCACCAACTGTGGGCCCTCACTATCCATATCCCAAGACTCTGC	7140
QY	7141	cccttcacccccaatgagatctcgggttggaagaccacaatgacctaggttaggaatgc	7200
Db	7141	CCTTTACTTCCTTATAGATCTGGGTGGAGGACCAATGCGCTTAGGCTGACGAATGC	7200
QY	7201	tgatgtccctcacactggaatctccctggagctggaagccccaagtctcaactggtctc	7260
Db	7201	TGATGTCTCTACACTGGATGTCCTGGACGTGGGTGAGCCCCAGTGTCCACTGTGTTCT	7260
QY	7261	gcccctagaacctataaggcgccctcccaccatcccaccagacttttggttctcttagag	7320
Db	7261	GCCCTAGACCTTATAGGGGCCCTCCCCCATCCCCACAGATTTTGGTTCTGTAGAG	7320
QY	7321	gtctttagcccaagccaagcggtgttgagaagagtgtgtttataacttaatgcaaaagac	7380
Db	7321	GTCTTAGCCCAAGCCAGCGGTGTGAGGACATGGTGTATTATTAACCTTAAGCAAAAGC	7380
QY	7381	tttcccccagaagacgtctcaaaatcttcccccccaccccccaacacacatcacacaca	7440
Db	7381	TTTCCCCAGACAGTAAATTTTCCCTCCCCACCCCAACACACATACACACACA	7440
QY	7441	ctctgtcagaagaacactgycctcgtagcaacccctccctctctacagcccaggtgtgtcagaag	7500
Db	7441	CTCTGCAGAGACACCTGCGCTGACCAACCTCCCTCTACAGCCAGGTGTTCAGAAAG	7500
QY	7501	gaagtcctaaaggagacttgagagaagggcccaagtgctaaagggccccagaagaagccgaagcc	7560
Db	7501	GAGTCCTAAGGGAGATGAGAGAGGCGCCCAAGGCTTAAAGGCCCCCAAGAACCCAGAGCC	7560
QY	7561	ttgagctcggggggggcgagaggtctggagaagcagaaactggaatgacctgagcaact	7620
Db	7561	TTGAGCTCGGGGGGGGCGAAGGTTTGGAGGACGAACTGGATCCTTGAGCAAACT	7620
QY	7621	gggacctaatcctaattaggtgttcgccagccccaagaagcctgtggccaattaaaccttcaa	7680
Db	7621	GGGCGCTATCTAATTAGGCGTTCGCCAGCCCAAGAGCGCTGGGCAATTAACTCTTCAA	7680
QY	7681	gtgcctcaactggaagaactcagaaggagagatcagctgttactctctccatggtccccaaga	7740
Db	7681	GTGCGCTACACTGAAMACACAGGGAGAGATCAGTTTACTCTCTCATAGTGTCCCAAGA	7740
QY	7741	gggtctctgggtgcccctgtgctcatctcccaactcagaaggtttgtctctccctggagtc	7800
Db	7741	GGGTCTCTGGGTGCCCTGTGCTATTCCCACTCAGAGGTTTGTGTTCTCTGGCAATC	7800
QY	7801	taaacctcagttgtctctctgtgctgagcagctgccccgttggagagctcttgtaattga	7860
Db	7801	TAAACCTCAGTTGTCGTCTGTGCTGGCAAGCTGCCCGCTGGAAGCTTGTAATGTA	7860
QY	7861	caaggacatcagaagtgagactggaatgaggagatacatataggatgtagccaatbagcactc	7920
Db	7861	CMAAGCATCAGAGGTGACATGCGAATGGGAGTACATAGGAGTGAAGCAAAATAGCACTTC	7920
QY	7921	aagggtgggtgagataataaagaactgtgcacccctgacactaagaagagcctaactatgat	7980
Db	7921	AAGGTGGGTGATATACAAATAAAGCTTTGTACCCCTGACCTTAGAAAGCTACTATATAT	7980
QY	7981	gatcacaaattgtttgacatcactctgtggagaatgtagtgaagacctaagctcaaaaacacacac	8040
Db	7981	GATCACAAATTGTTGACATCACTCTGGGACATGTAGTGAAGACCTTAGCTCAAAACACACAC	8040
QY	8041	agtagctttaagagtcaggtctgtgacttaatactggaactcagggccctaagtgtgtcg	8100

Db	8041	AGTAGCTTTAAGAGTCAGCTGTGACTTAACTCGAATCTCGAGCCCTAATAGGGCTGG	81001
QY	8101	gtatgtctgcgcccaacacccctgttttagttagatctctgtgcatactccacccacgttgg	81606
Db	8101	GTATGTCTGCCCTACACCCCTGTTTAGTAGATCTCTGCGTAATCTCACCCACAGCTGG	81606
QY	8161	gtgggctgtctcttccctctgaggcgaggaatgtgtctctccatcagaatagagccgg	82200
Db	8161	GTGGGGCTGCTCTCTCCCTTGAGGGCAGGAATGTGTCTTCCATCAAGATAGGACCGG	82200
QY	8221	tgtrtagaagaacatctgtctgtcgtgttctctgaatatataatgaacgtaactctacagc	82800
Db	8221	TGTAGTAGAAGCAACTGCGCTGCGCTGTGTTCTTGGAATAATTAAGAGCAATCTATACAGC	82800
QY	8281	ctgggttagtagaactaaacaaggggtgtgggggctgtctctgcgaacacgagataggtatag	83400
Db	8281	CTGGGTTAGTAGACTAAACAGGGGCTGTGCTCTGCGAAGAACGATAGGCTATAGG	83400
QY	8341	agccactgcagcctagatttaacacacttgggtgttctgtcactagagccattctcaacaagc	84000
Db	8341	AGCCACTGCAGCCTAGATTAAACACTTGGGTGTCTGTCTAGTACAGCCATTCTGCACAAAGC	84000
QY	8401	agttcctcaagacactcgggagacactgtttgcagcattaatgtccagcattaatgtcagcatc	84600
Db	8401	AGTCTCTAAGACACTGGGAGCACTGTTCGACGATTAATCCAGCATTTAATGCCAGCAATT	84600
QY	8461	aggggaggcagagagcagaagaatctctcgaattcaagaagccatcttgatattacaataag	85200
Db	8461	AGGGGAGGCGAGAGGCGAMAGATCTCTGTAGTCAAGGCCATCTGAAATTAACATAAG	85200
QY	8521	agctccagcgccagccaggggtgtgcgaatlaaacacttgtctccaaataaacaagcatcttag	85800
Db	8521	AGCTCCAGGCGCACCGAGGGTGGCCAGTAAACCTGTGTCTCAAAAAACAAGATCTTAG	85800
QY	8581	tgaccaagctctgtctcaaccgccagtgtaaccaaggaaccccccacccgagtgcaagttagcc	86400
Db	8581	TGACCAAGCTCTGTCTCAACCCCAAGTACCAACGAGACCCCCCACCCAGCATGTACAGCAAGCC	86400
QY	8641	ggcgttggggccctcggaggagacacgtatgctgtcgtgtgtctcacacaagatctcgaag	87000
Db	8641	GGC GTTGGGGCCTTGAGAGACCACTAGTGTGGTGGCTCTACACCACTAGTCTCAAGG	87000
QY	8701	attctccttccaaagccaagtiacacagatccgctacccggctggaggacagcgttgcagtga	87606
Db	8701	ATTCTCTCTTCCAAGCCAAAGTAACTCGCTACCGCTGAGAGCAAGCGTGAAGTGA	87606
QY	8761	aggtgtcccggtcccgcccggaacccgccttgaaaccccgcccccagcatctgaactctccct	88200
Db	8761	AGGTGCCCGGCTCCGCCGCCGAGACCCGCCCTTGACCCCGCCCCGATCTGACTCCTCCT	88200
QY	8821	cacccgttcaggtgtgtatgtatgactgacgacaaacagacactctgcgtctcggggcgctgaa	88800
Db	8821	CACCGTTCAGGTGTGTATGTATGACTGACGACAAACAGACCTCTGCCTCTCGGGCGCTGAA	88800
QY	8881	gccccgagaccgttactctgtctccaaagtgtgtgttaaccattcggatcctatgtgtcgaa	89400
Db	8881	GCCCCGAGACGTTACTCTGTCTCCAAATGTGGTGTGAACCATTCGGGATCTATGGTCAA	89400
QY	8941	aaagcgaggaaacttggagcgagtgtgagacaccccaaccgcgtgtcctccacccctcgaaatgg	90000
Db	8941	AAAGCGGGAATCTGAGAGCGAGTGTGAGCCACCCACCGCTGCTCCACCCCTCGAAGTGG	90000
QY	9001	tgagcactcttccaaagggctgtcgtgcgcctatgtaattcccccaatcactctgttctctccc	90600
Db	9001	TGAGCACTCTTCCAAGGAGCTGGCTGGGCCAATGAATCCCCAAATCACTCTGTCTCTCCC	90600
QY	9061	ccccacccttttttttggagaaagcgcttctcaagtagcagcatgtcgtgcctaaatcagtatg	91200
Db	9061	CCACACCCTTTTTTTTGAAGACAGGCTCTTCAAGTAGAGGCAATGCTGGCCTTAATTCAGTAG	91200
QY	9121	tagtcaagatgaactcagactcgtctgtcttcttctcaccattagagacaatgycgaat	91800

Dh 9121 TAGTCAGGATGACCTGAGCTCCTGCTTTTGTCTCCACTTAGACACAATGCCACT 9180
Qy 9181 ggccatcaacaccccttggagactagcaatgagatctatttagctctgcatcttgtaga 9240
Dh 9181 GGCATCATCCACCTTTGGGAGACTAGCCATGAGTATTTAGCTGTGCAATTTGGTGACA 9240
Qy 9241 gatgagatcaacacggttgacctcttgtaagagaacatgaaagagctgttttaacccc 9300
Dh 9241 GATGAGATACACAGTGTGACCTCTTGTAAAGAACTGAAGACAGCGCTGTTTTAAACCCC 9300
Qy 9301 aatactcaagctctctagaggttaaccttatataaataagagctcttaacagcagta 9360
Dh 9301 AATTTCTAGGCTCTCTAGAGTTAACTTTATATAAATAGACATTTACAGCACTTA 9360
Qy 9361 tcacatggtcccaacagaaaccttttgcacacaaacta tagaccacagtgctgtgtctac 9420
Dh 9361 TCACATGAGTCCACAGAACCTTTGTCAACAACCTTAGACCAACATGGCTGTGCTAC 9420
Qy 9421 caataagaggtctctactgtctgtgcccacccctcaaccccttaaaaggttaacctaggcagc 9480
Dh 9421 CACATAAGGGTCTCTACTGTGCGCCACCCCTCCAAACCTTAAAGGTAACTAGCGCAGC 9480
Qy 9481 cttaatacttgcaactctctactactcaagctcttgtaagtctagaagaacagagcttaacc 9540
Dh 9481 CTTAATATTGGCAATCTCTACTCTCACTGACCTCTTGAATGCTCAGAAACCAAGGCAATTAC 9540
Qy 9541 caaagttctctctctgtgtgtcccttctcttaagtgtagagggccctaaagatgactctctt 9600
Dh 9541 CAAGTTCTCTCTCTGTGTGCGCTCTTCTTAAGGGGAGGCGCTAAAGATGACTCTCTT 9600
Qy 9601 gtcttgaaagactctcgagcccaatgagctgtcaactcttaatatgaatatattgtacaa 9660
Dh 9601 GTCTTGAAAGACTCTCCAGCCCAATGATGTGCACTCTCTAATATGAATATATGCTAA 9660
Qy 9661 aatgtctgscctgaagtttccccacactgtctaggttgagcagcaagtcggttccaagaac 9720
Dh 9661 AATGTCTGSCCTGAGTTTCCCCACCTGTCTAGGTTAGGACACAGCGGTCTCAAGACAC 9720
Qy 9721 ttaatacttgacagcagataaagaagaagctccatcccccacccgctctccctcgctgc 9780
Dh 9721 TTATATTTCGAGGCGATTAAGAAAGTCCCATCCCAACCCGCTCTCTCCGCTCC 9780
Qy 9781 ctaagacagaatactctctactcaactgaaactctctcgacagcagctatgctcaacttaa 9840
Dh 9781 CTAAAGACAGATACTCTTCACTGAAACTGAACCTCTCGACAGCAATATGCTCACTTTAA 9840
Qy 9841 tgaatgtaataataatgaggaactgagctccgaagagatctccgtgaggaagaaggtcaaa 9900
Dh 9841 TGATGATGAATAATAGGGGAACTGAGGCTCCGAGAGATTCCTGAGGAGAGGCTCAAA 9900
Qy 9901 accagctccagaagactctctcagcccccatccgagctctcagagttcttgaggttgag 9960
Dh 9901 ACCAGCTCCAGAAAGCTCTCTCAACCCCAACCCGCTCTCCAGGTTCTGAGGCTTGGCG 9960
Qy 9961 gagtgaacacagctggaggggctgtgagccttggaagcttggccttgctgtgtgcccagc 10020
Dh 9961 GAGTGAACACAGCTGGAGGGGCTGAGAGCTTGAGAGCTTGGCCTTGCTGCGGCCAGC 10020
Qy 10021 acctgagatcttgacagcggagcagcagcagcagcagcagcagcagcagcagcagcagc 10080
Dh 10021 ACCTGAGATCTTGACAGGGAGCAGCAGGCGGCTGCTCCGCGCAGAGACTGAAGAAG 10080
Qy 10081 ccgggggttagaggttgaggaaggaagcaagcagggcgtgtgggggagcgaagcttgccagag 10140
Dh 10081 CCGGGGTTAGGCTTGAAGGAGGTAAAGCAGGGGCTGTGGGGCGCGAAGCTTGTGCCAGG 10140
Qy 10141 cctgtcagcagagtcctccagtttattatgagcgtgagcagatgctctatccgctgagc 10200
Dh 10141 CCTGTACGAGAGTCCCAAGTTTATTTATGCGTGAAGCCGAGATGCTTATCCGCTGAGC 10200
Qy 10201 tgcctggagagatgctgagcgtgagggatgagcccaagggctgtgctcccaactcagctcc 10260
Dh 10201 TGCTGGAGGATGCTGCGCTGGGGATTTGAGACCAAGGGCTGCTCCACTCAGTCCATC 10260

Qy 10261 cagcccatcagatgacacacccgtgcatctctgaagcttatcttggaacccgcccctg 10320
Dh 10261 CAGCCCATCTCAGTGTACACACCCGTGCTCTCTGAGCTTATCTTGGGAAACCCGCTTG 10320
Qy 10321 tctgtgtgtctgtctctatctctgtcatctcaactcccaagagcctttttatgctt 10380
Dh 10321 TTCTGTGCTGTGTCTCTCTATTTCTGTGATCTACTTCCAGAGCCCTTTTATGCTT 10380
Qy 10381 ttaataactagcttttaaaatgtctttgataatgtgtgtgctctgtgaggtgc 10440
Dh 10381 TTATATATAGTGTGTTTAAATGCTTTGTATATATGTGTGCTGTGTGAGCTGTG 10440
Qy 10441 gtgcacaacacacagctgaggttgaagaaactgtgtgagtagggctctccacagctg 10500
Dh 10441 GTGCCACACACACAGCTGAAGTTAGAGAACTTTGTTGATGAGCTCTTCCACCATGT 10500
Qy 10501 gggactagggctgtgcgaagaagaacttaactgatactctgcagagccctcaacccctca 10560
Dh 10501 GGGACTAGGGCTGGCGAAGAGCAATTAAGTATCTGCGACGCCCTCAACCCCTCA 10560
Qy 10561 ctcccatcctgtttgagatgataagtaatcgaagtaaatcgtgtgcttaattcg 10620
Dh 10561 CTTCCTCATCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 10620
Qy 10621 tagctatcctgtcctcagccctcagagagtgctgtctcaaaagtttgaggaggggctcc 10680
Dh 10621 TAGCTATCTGCTCTCAGCCCTACCAAGTGTGTGTCTACCAAGTTGTGGAGGGGCTCTCC 10680
Qy 10681 tcccagtgctgtgggttacacagctcccaagatctcagctctcaggtcttgcttgtagtt 10740
Dh 10681 TCCAGTGTCTGTGGGTTACACAGATCTCCAGATCTCTGCTTCTTGTGCTTATGTT 10740
Qy 10741 tgcctctgtcttgcgtgtgtcctcagagctcagagctcagagctcagagctcagagct 10800
Dh 10741 TGCCCTGT 10800
Qy 10801 ctcttgacagaaatagctgttttaactcccatgatttgacccctctctgtctgtct 10860
Dh 10801 CTCTTGACAGAAATAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10860
Qy 10861 ccacgcgctgtgagctgcaactctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 10920
Dh 10861 CCATGCGCTGT 10920
Qy 10921 caacttcccaagcgaagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 10980
Dh 10921 CAACCTTCCCAAGCGAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10980
Qy 10981 gctgtgcgagcccccacacatgctgcctcaattctttaagagcccgagcgagcgagcg 11040
Dh 10981 GCTGTGCGGCGCCCAACACTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 11040
Qy 11041 ggggtgtcagcagcggggcgagcagcagcagcagcagcagcagcagcagcagcagcagc 11100
Dh 11041 GGGGTGTGCAAGCCGCGGGCGGCGAGCCAGCTGGGCGCGGTGTGCGCGGCGAGCTAAGC 11100
Qy 11101 agttccctggtgtgtcagaagcagcagcagcagcagcagcagcagcagcagcagcagc 11160
Dh 11101 AGTTCTCTGCTGT 11160
Qy 11161 accagtggt 11220
Dh 11161 ACCAGTGGGCTGT 11220
Qy 11221 gggagagctgt 11280
Dh 11221 GGGAGGCTTGT 11280
Qy 11281 gacaaacacgcaactctcttccaaagcagagcagagagagagagagagagagagagagag 11340
Dh 11281 CACAACACGCACTCTTCTTCCAAAGCAGAGCAGAGGAGGATCTGCTGCGGCGAGAGC 11340

[illegible]

RESULT	2
ID	V27145
AC	V27145 standard: DNA; 6663 BP.
DT	02-OCT-1998 (first entry)
DE	Nucleotide sequence of Murine NR6.
KM	Haemopoietin receptor; cell proliferation; cell differentiation;
KM	cancer; cell survival; therapeutic; neuronal proliferation; drug;
KM	screening; ss; Mouse.
OS	Mus sp.
EH	Key
FT	Location/Qualifiers
FT	1182..1744
FT	/*tag= a
FT	/product= "Murine NR6"
FT	/note= "No start or stop codon given"
PD	W09811225-A2.
PD	19-MAR-1998.
PF	11-SEP-1997; G02479.
PR	11-SEP-1996; AU-002246.
PA	(AMRA-) AMRAD OPERATIONS PTY LTD.
PA	(DZIE/) DZIEGLEWSKA H E.
PI	Alexander W, Fedrli L, Farley A, Hilton DJ, Kikuchi Y,
PI	Kojima T, Maeda M, Nash A, Nicola NA, Rakar S, Willson T,
PI	Zhang J;
DR	WPI; 98-260970/23.
DR	P-PDB; W55016.
PT	New isolated haemopoietin receptor - used for developing products
PT	for modulating proliferation, differentiation and survival of cells,
PT	e.g. neuronal cells
PS	Claim 8; Page 108-114; 182pp; English.
CC	The NR6 gene encodes a novel Haemopoietin receptor (HR). Interaction
CC	between the novel HR and a ligand facilitates proliferation,
CC	differentiation and survival of a wide variety of cells. The HR and it's
CC	products can be used for modulating the activity of the receptors e.g. to
CC	regulate development, maintenance or regeneration in an array of
CC	different cells and tissues in vitro and in vivo. They can be present in

CC therapeutics used for modulating neuronal proliferation, differentiation
CC and survival. The products can also be used for detection and diagnosis
CC e.g. for cancers or predisposition to cancers, or for drug screening.
SQ Sequence 6663 BP; 1462 A; 1852 C; 1715 G; 1634 T;

Query Match	55.5%	Score	6561.6	DB 1	Length	6663
Best Local Similarity	99.6%	Pred. NO.	0			
Matches 6628	Conservative	0	Mismatches	19	Indels	5
					Gaps	5

OY	5186	tcgacgcgtgaagccagggaggtt-ccaagttccaagccagttg-cttcaggaatgaagc	5243
Db	12	tgagacctgacagcaggagagttcccaagtttcagacacatgtgtttcaggaatgaac	71
OY	5244	ccgtlcaagaaagaaagaaatcaaaagagcagaagaaatgcttataagctgtgaacagc	5303
Db	72	cctgcacaagaaaaaagaaagaaatgaagacacagaaatgctttatagctgtgaacagc	131
OY	5304	ttggttggttaaggggcactgtgcctcaatcaatgaatgacctcagccccccctlaagaaic	5363
Db	132	ttggttggttaaggggcactgtgcctcaatcaatgaatgacctcagccccccctlaagaaic	191
OY	5364	catgttgaggaggagaaagcaaac-ccagctgtgcgcccaatacatgtgtcccaatgt	5422
Db	192	catgttgaggaggagaaagcaaacctccagctgtgcacctccatcatgtgtcccaatgtg	251
OY	5423	caacacacacaggggagacataatcaatataagatgtatttcctagattgaagagca	5488
Db	252	cacacacacacaggagacatattatattatagatgtatttcctagattgaagagca	311
OY	5483	ttatgacctagtlttaaaatttattatgattttatgaataatactggttttatt	5542
Db	312	tttatgacctagtgtttttaaatttttattgattttatgaataatactggttttatt	371
OY	5543	ggttggttggttggttggtttgttttattatgagacaggtctctctgtgtagctgct	5602
Db	372	gggttggttggttggttggtttgttttattgagacagggcttctctctgtgtagctgct	431
OY	5603	gtcccttggaactcaactctgtgtagacacagggcgtgagcttgaaactagaatccgcctgct	5665
Db	432	gtcccttggaactcaactctgtgtagacacagggcgttgaactagaatccgcctgctgt	491
OY	5663	gcttcccaagtgcttaagattaaagtggtgcactgccaatcaagcaaatgtaactttaa	5722
Db	492	gcttcccaagtgcttaagattaaagtggtgcactgccaatcaagcaaatgtaactttaa	551
OY	5723	cccccaagatttgggagagccagagccagacaaatggtgtgaattccaggtcagccaagatac	5782
Db	552	cccccaagatttgggagagccagagccagacaaatggtgaattccaggtcagccaagatac	611
OY	5783	agagtgagaccctacttcttaccctcccccaccacaccccaaatgtaattgtgtgctgt	5842
Db	612	agagtgagaccctacttcttaccctcccccaccacaccccaaatgtaattgtgtgctgt	671
OY	5843	gtatgtacatgtgtgtgtagcagcagtaaaatgtlccaagagcaactgtgaagctctcc	5902
Db	672	gtatgtacatgtgtgtgtagcagcagtaaaatgtlccaagagcaactgtgaagctctcc	731
OY	5903	gttcaacagcttaagtcctgaattcaacaactaaggtctccaggtctgagccaagctctctt	5962
Db	732	gttcaacagcttaagtcctgaattcaacaactaaggtctctccaggtctgagccaagctctt	791
OY	5963	atgtaacagcaattcaacagcgcctggagattgaacgataatataattttttagataaagt	6022
Db	792	atgtaacagcaattcaacagcgcctggagattgaacgataatataattttttagataaagt	851
OY	6023	ctcttgtagcttaagcttaaggtctcaaaacttgaactcccaaggtlcaacttgagctgtgt	6082
Db	852	ctcttgtagcttaagcttaaggtctcaaaacttgaactcccaaggtlcaacttgagctgtgt	911
OY	6083	actcttgcttccacccccaagtggtgtgaatgatacccaagcagcactctctcggggagagg	6142
Db	912	actcttgcttccacccccaagtggtgtgaatgatacccaagcagcactctctcggggagagg	971

[illegible][illegible]

Db 3132 GTGGGGGCGTGGTGGAAAGCAGATAGGCTATAGAGAGCCAGCCAGCTAGATAC 3191
Qy 8362 accactgggtgtctgtctctctctctctctctctctctctctctctctctctct 8421
Db 3192 ACCACTGGGTGTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3251
Qy 8422 ttgtccagcatltaatgccaagcatttaatgccaagcatttaatgccaagcattta 8481
Db 3252 TGTTCAGCATTTATATGCGACATTTATGCGACATTTATGCGACATTTATGCGAC 3311
Qy 8482 atctctcgaagtcgaagcaccatctgaalttaataaagagtcgaagcaccaggtg 8541
Db 3312 ATCTCTGAGTTCAAGGCGCATCTGAATTTACATTAAGAGCTCCAGCGACAGCG 3371
Qy 8542 cgcagtaaaccttctcgaataaagaagccttcttaagtcgaagccttctctcacc 8601
Db 3372 CGCAGTAAACCTTGTCTCAAAAAAAGCATCTTACTGACAGCGCTGCTCCACCC 3431
Qy 8602 cagtaaccagagaccaccaccagcgtgacagtcgagccgcttgggggctggagacc 8661
Db 3432 CAGTACACAGAGACCCGCCACCGACGTCAGTGAGCGCGCTTGGGGGCTGGAGACC 3491
Qy 8662 agctgagtgctgctggtgtctcaaccagcgtctcaaggaattctctctcgaagcca 8721
Db 3492 AGCTGAGTGTGGGTGGGTCTCACCACAGCTCTCAAGATTTCTCTCTCAAGCCAA 3551
Qy 8722 accagatccgctaccagcgtgtggagagcagcgtgtgagagtgccgctccgcccga 8781
Db 3552 ACCAGATCCGCTACCGCTGGAGAGACGTCGAGTGAAGGTGCCGCCGCCGCCGA 3611
Qy 8782 cccgcccctgaccccccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 8841
Db 3612 CCCGCCCTGACCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCC 3671
Qy 8842 cgtcagcaaccagagcctctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 8901
Db 3672 CGTCAGCAACACAGACCTCTCTGCGCTGCGGGGCTGAAGCCCGGACCGTTACTCT 3731
Qy 8902 ccaagtgctgtgttaaccatctcgtgagctctgaagtcgaataaagcgaggaatct 8961
Db 3732 CCAAGTGGGTGTGAACCATTCGAGATCTATGGGTGCAAAAAGCGGGGATCTGAG 3791
Qy 8962 gtggagcaaccacagcgtctcgaacccctcgaagtggtgagcaccctcgaaggtct 9021
Db 3792 GTGAGCCACCCACCGCTGCTCCACCCCTGGAAGTGGTGAACCTCTCAGAGGCT 3851
Qy 9022 ctggcccaatggaatcccaatccatccatctctcctcccccaccccttttttgaag 9081
Db 3852 CTGGCCCATGGAATCCCAATCCATCCCTGTCCTCCCCCAACCTTTTGTGAGAC 3911
Qy 9082 cgtctcaggaagcagcgtctgaccttaaatcagatgtagtcaggaatgacctcga 9141
Db 3912 CGTCTTCAAGTGAAGGCACTGCGCTTAATTAATGATGATGATGATGATGATGAT 3971
Qy 9142 tccggtgtcttctctcgaattagagcaatgagcagtgagcaccacacacacacacac 9201
Db 3972 TCCGAGTCTTTTGTCTCACTTGAAGCAATGGCGAGTGGCCATCACACCTTTGG 4031
Qy 9202 actagcaatgagctatctagcctgtcatctgtgtgagagatgaggaatgaacagtgt 9261
Db 4032 ACTAGCCATGAGCTATATTAGCCGTGATTTGGAGAGATGAGATGAGATGAGATG 4091
Qy 9262 cctcttgaagaagcgaagaagcgtgttcttaaccccaatctcgaagcctctcga 9321
Db 4092 CTCTTGAAGAGACTGAGAGAGAGGCTGTTTAACTTAACTTAACTTAACTTAACT 4151
Qy 9322 gttacttataataatagagctatctacagccagcttatacaatgagcccaagacct 9381
Db 4152 GTTAACTTATAATAAAGAGACTATTACAGCCAGTTATACAGAGGTCACACAAACCT 4211
Qy 9382 ttgtcacacacacatagacacagctcgtgtcctacacataaggtctctcgtct 9441

Db 4212 TTGTCAACAACCTATAGACCAAGTGGCTGTGCTTACCATATAAGGCTCTCTACT 4271
Qy 9442 ggcacacccctccaaaccttaaaaggtlaacctaggaagccttaataattgcaatcc 9501
Db 4272 GGCACACCCCTCCAAACCCCTTAAAGGTAACTTAGGACAGCTTAATATTTGCAAT 4331
Qy 9502 acctagcctcttgaatgtctcgaagaacagagcatttaacccaagttctctctcgg 9561
Db 4332 ACTCAGCCTCTTTGAATGCTCAGAAACAGGCAATTAACCAAGTTCTCTCTGGG 4391
Qy 9562 ccttcttaagggtggagagccttaaaagatgactccttctcctgaagcctccagc 9621
Db 4392 CCTTCTTAAGTGGAGGAGGCTTAAAGATGACTTCTTGTCTTAAGACTCTCGAGC 4451
Qy 9622 catgtatctgaactctctaaatgaataatltgataaataatgtctgagctctcc 9681
Db 4452 CATGATCTGCACTCTCTAATATGAATATGATATGAATATGATGATGATGAT 4511
Qy 9682 caactgtcaggtttaggcagcagtcggtccaaagccttaataattgcaagcag 9741
Db 4512 CACCTGTAGGTTTAGGAGCAGACAGTGGTCCAGACACTTCTATTATTTGAGCA 4571
Qy 9742 aagaagaagctcccaatcccccacccgctcctcgtctccctgaagcagaatctct 9801
Db 4572 AAGAAGAACTCCATCCCCACCCGCTTCTCCGCTCTTAAGCAGATCTCTTACA 4631
Qy 9802 ctgaaactgaactctcgaagcagcatatgtctactttaaataatgataatgagga 9861
Db 4632 CTGAACCTGAACCTCTGAGAGAGCATATGCTCACTTATATGATGATGATGAT 4691
Qy 9862 actgagctccgaagagatctcgtgagagagaggtcaaaacagctccagaagctcc 9921
Db 4692 ACTGAGCTCCGAGAGATCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4751
Qy 9922 agcccccacccgggctctcgaaggtctcggaggttggcggaatgagcaagctgg 9981
Db 4752 AGCCCCATCCGGGCTCTCCAGGTTCTGGCTTGGCGGAGTGAACAGCTGGAGAG 4811
Qy 9982 gcttgaagccttggagccttggccttctgtcccaagcactcgaatctctcgaag 10041
Db 4812 GCTGAGCCTGGAGACTTGGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTT 4871
Qy 10042 gccaaagcagcgtctcgtccgcgcgaagagactgaagaacgggggtgaggttga 10101
Db 4872 GCCAAGCAGGCGCTCGTCCGCCGAGAGACTGAAGAAGCGGGGTGAGGTGAGGA 4931
Qy 10102 ggttaagcagggcgttggggggcggaagctgtgccaagcctgtgaagcagctcc 10161
Db 4932 GTTAAGCAGGGGCTGTGGGGGCGGAAGCTTGTGCCAGGGCTGTACGAGATCC 4991
Qy 10162 ttattatgctgtgagcagatgtctctatccgctgtgctgtggtgaggtgctcgtc 10221
Db 4992 TTATTATGGGCTGAGGCGGATGCTTATCCGCTGAGGCTCTGAGGAGATGCTG 5051
Qy 10222 ggggacttgaaccaaagcgtgtctcccaactagtcctcgaagcccaatcatal 10281
Db 5052 GGGGATTTGAGACCAAGGCTGCTCCCATCAGTCCAGCCACACTGCATGTCACAC 5111
Qy 10282 cgtgactctctgagcgtctactcttgggaacccgacctgtctgtcgtcgtctcat 10341
Db 5112 CGTGATCTCTGAGGCTTATCTTGGGAACCGCCCTGTTCTGTGTGTCTCTAT 5171
Qy 10342 ttcttcatcacttcccaagccttttttcttcttcttcttcttcttcttcttctt 10401
Db 5172 TTCTGCTATCACTTCCCAAGCCTTTTATGCTTATTAATTAATTAATTAATTAAT 5231
Qy 10402 aatgctctgtataatgtgtgtcctcgtgtgagcgtgtgtgacacacacacacac 10461
Db 5232 AATGCTTTGTATAATGTGTGCTCTGAGCGTGTGACGTCGACACACACACAG 5291
Qy 10462 ggttagaagaacttgtttagagagctcctccacacagtgagagctagggcggagaa 10521
Db 5292 GGTAGGAACCTTTGTGTAGTAGCTCTTCCACCATGTGAGATGAGGCTGAGCAAG 5351

[illegible][illegible]

	CC	derivatives can be used for modulating the activity of the receptors e.g.
	CC	to regulate development, maintenance or regeneration in an array of
	CC	different cells and tissues in vitro and in vivo. They can be present in
	CC	therapeutics used for modulating neuronal proliferation, differentiation
	CC	and survival. The products can also be used for detection and diagnosis,
	CC	e.g. for cancers or predisposition to cancers, or for drug screening.
	SQ	Sequence 938 bp; 243 A; 245 C; 272 G; 178 T;

Query Match	6.2%	Score 737.2;	DB 1;	Length 938;
Best Local Similarity	96.2%;	Pred. No. 1.2e-162;		
Matches 777;	Conservative 0;	Mismatches 28;	Indels 3;	Gaps 2
OY 10983 tggcgcgcccccaaacactgcgcgtccatcttcttagagcgccggggcccggcgcgcg				11042
Db 79 TGGAGCCACACCACCGCTGCCCTCCACCCCCTTCGAAATGACGCCCGCGCGCCGGCGGCGG				138
OY 11043 gttgtccaggccgg				11102
Db 139 GTGTCTCGAACCCCGGGGGCGCGCAGCCCGACTGGGGCCCCGGTCCGCGAGCTCAAGCG				198
OY 11103 ttccctcggttggtctaagaagacgcatactgtcgaacctagtcttcgcgtgtacac				11162
Db 139 TTCCTGGGTGGCTCAAGAAGCACCCATACTCTGGAACCTTAGTTCCGCTTGTAACGAC				258
OY 11163 cagttggttgtcttgaatgcagaagttcacacaagaccgaaacaggtlagaagtttggg				11222
Db 259 CAGTGGCGCTGCTTGGATGCAGAAAGTCAACAAGACCCGAAACAGTAGAAGATTGGGG				318
OY 11223 gaggtttgcgttggatgcagaagttcacacaagaccgaaacaggtlagaagtttggg				11282
Db 319 GAGGCTTCCTGTGGGGGGAAGAGACCAAGAGAGAGAGACCCCGGTGAGACGCTCTCA				378
OY 11283 caacaccgcactctcttctccaagacagagagaggtatcctgccctcgsgcaacaggg				11342
Db 379 CAACACCCACTCTTCTTCCAAGCACAGAGAGAGAGGATCTGCTCGGGCAACAGCGG				438
OY 11343 gtgcggcgagagatlaaggggtctgtgtgaattggggcttacagcagtttagatgaagcc				11402
Db 439 GTGCGCGAGAGATGAAGGGGGTCTGGGTGAGTGGGGCTTACAGCAGTCTAGATGAGGCC				498
OY 11403 ttccctctcttgcgtgtgtgtctcaaggatctcttagtgcatacttaaccactycaaa				11462
Db 499 TTTCCTCTCTTGGGTGTGCTCAAAGGAGATCTTATGTGCTCATTCACCACTGCAAA				558
OY 11463 gagccccaggtttaactgcatcatcaaatgttcgaaggtccaaggtctaagtgtgctct				11522
Db 559 GAGCCCCAGGTTTACTGTCATCATCAATATTCCTGAAGGGTCCAGGCTTAAGTGGCCCT				618
OY 11523 ttctgtccctcaaggtccctgcgcgcgtctaaactcaagatcaagcaccctccgtgtgtca				11582
Db 619 TTTGCGCCCTCAGGCTCCCGCGGCTCAAACTTAAGATGAGGCCATCTCTGCTGGGTCA				678
OY 11583 gaactggaaggtcactgaatlttggagccctctylac - atctggycacaagaagaact				11641
Db 679 GACCTGGAGGCTCACTGAATTTGAGGCCCTCTGACTACTGTGGGCACAAAGAAGACT				738
OY 11642 acca--gaggtctgggcaaatgagatccccacaacacagctttgtccacatgatgtica				11699
Db 739 ACCATGAGGCTGGGGGCACAAAGAGACTCCCAACACAGCTTTGTGTCACATGATGTC				798
OY 11700 caacttgatatccccaaggttgggttaggttgggttatitgaagggccctccaagaagctc				11759
Db 799 CACTTGATATACCCCAAGTGGGTAAAGTTGGGGTATGTGAGGGCCCTCCACACATCTC				858
OY 11760 ttaaataaataaagaaggttgcagtt 11787				
Db 859 TTTAAATAAATAAAGAGATTGTCAGG 886				

RESULT	5
ID	V27140 standard; CDNA; 1629 BP.
	727140

AC V27140:
 DT 29-SEP-1998 (first entry)
 DE Novel haemopoietin receptor NR6.1 gene.
 KW Haemopoietin receptor; cell proliferation; cell differentiation; cancer;
 KM cell survival; therapeutic; neuronal proliferation; drug screening; ss;
 OS Mouse.
 FT Key
 FT CDS
 Location/Qualifiers
 113..1355
 /tag= a
 /product= "Haemopoietin receptor NR6.1"
 WT
 PN WO9811225-A2.
 PD 19-MAR-1998.
 PF 11-SEP-1997; G02479.
 PR 11-SEP-1996; AU-002246.
 PA (AMRA-) AMRAD OPERATIONS PTY LTD.
 PI (DZIE/) DZIEGLEWSKA H E.
 PI Alexander W, Fabrl L, Farley A, Hilton DJ, Kikuchi Y,
 PI Kojima T, Maeda M, Nash A, Nicola NA, Rakar S, Willson T,
 PI Zhang J;
 PI WPI: 98-260970/23.
 DR P-PSDB: W55011.
 PT New isolated haemopoietin receptor - used for developing products
 PT for modulating proliferation, differentiation and survival of cells,
 PT e.g. neuronal cells
 PS Claim 4: Page 77-81; 182pp; English.
 CC The haemopoietin receptor (HR) NR6.1 is a form of the novel HR NR6.
 CC Interaction between the novel HR and a ligand facilitates proliferation,
 CC differentiation and survival of a wide variety of cells. The HR and its
 CC derivatives can be used for modulating the activity of the receptors e.g.
 CC to regulate development, maintenance or regeneration in an array of
 CC different cells and tissues in vitro and in vivo. They can be present in
 CC therapeutics used for modulating neuronal proliferation, differentiation
 CC and survival. The products can also be used for detection and diagnosis,
 CC e.g. for cancers or predisposition to cancers, or for drug screening.
 SQ Sequence 1629 BP; 336 A; 541 C; 453 G; 299 T;

Query Match 2.4%; Score 280.4; DB 1; Length 1629;
 Best Local Similarity 99.6%; Pred. No. 6,9e-56;
 Matches 281; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6330 acacagctgtatcagccccagagccccaccccttcctcagctccctccgaagcta 6389
 |||||||
 DB 248 ACACAGCTGTATCAGCCCCAGAGCCCCACCCCTTCTCATCGCTCCCTGCAAGCTA 307
 |||||||
 QY 6390 cctgtctatacatgagacacacctgaggccacgcctgagggctctactgagacctta 6449
 |||||||
 DB 308 CCTGCTCTATACATGAGACACACCTGGGGCCACCGCTGAGGGGCTCTACTGAGCCCTCA 367
 |||||||
 QY 6450 atgtgtcgcgcgcctcctgagctgtccgcctccttaaacctccacccctggccttg 6509
 |||||||
 DB 368 ATGTGTGCGCGCTGCGCTCTGAGCTGTCCGGCTCTTAACACCTCCAGCCCTGG 427
 |||||||
 QY 6510 cccctgtctaaccttaatgtgtccagagcagctcagagagacaaatctgtgttcaagccc 6569
 |||||||
 DB 428 CCTGTGCTAACCTTATGTGGTCCAGGACAGTCAAGAAACATCTGTGTGTACGCC 487
 |||||||
 QY 6570 gagacgagcagcatctgtgctgtgctcctgctctatgttgcct 6611
 |||||||
 DB 488 GAGACGGCAGCATTTCTGTGGCTGTGCTCTGCTCTATGTGTGCT 529
 |||||||

RESULT 6
 V27141
 ID V27141 standard; CDNA; 1673 BP.
 AC V27141;
 DT 29-SEP-1998 (first entry)
 DE Novel haemopoietin receptor NR6.2 gene.
 KW Haemopoietin receptor; cell proliferation; cell differentiation; cancer;
 KM cell survival; therapeutic; neuronal proliferation; drug screening; ss;
 OS Mouse.
 FT Key
 FT CDS
 Location/Qualifiers
 122..1399
 /tag= a
 /product= "U4 protein"

FT Key
 FT CDS
 Location/Qualifiers
 1..1278
 /tag= a
 /product= "Haemopoietin receptor NR6.2"
 WT
 PN WO9811225-A2.
 PD 19-MAR-1998.
 PF 11-SEP-1997; G02479.
 PR 11-SEP-1996; AU-002246.
 PA (AMRA-) AMRAD OPERATIONS PTY LTD.
 PI (DZIE/) DZIEGLEWSKA H E.
 PI Alexander W, Fabrl L, Farley A, Hilton DJ, Kikuchi Y,
 PI Kojima T, Maeda M, Nash A, Nicola NA, Rakar S, Willson T,
 PI Zhang J;
 PI WPI: 98-260970/23.
 DR P-PSDB: W55012.
 PT New isolated haemopoietin receptor - used for developing products
 PT for modulating proliferation, differentiation and survival of cells,
 PT e.g. neuronal cells
 PS Claim 5: Page 84-87; 182pp; English.
 CC The haemopoietin receptor (HR) NR6.2 is a form of the novel HR NR6.
 CC Interaction between the novel HR and a ligand facilitates proliferation,
 CC differentiation and survival of a wide variety of cells. The HR and its
 CC derivatives can be used for modulating the activity of the receptors e.g.
 CC to regulate development, maintenance or regeneration in an array of
 CC different cells and tissues in vitro and in vivo. They can be present in
 CC therapeutics used for modulating neuronal proliferation, differentiation
 CC and survival. The products can also be used for detection and diagnosis,
 CC e.g. for cancers or predisposition to cancers, or for drug screening.
 SQ Sequence 1673 BP; 344 A; 550 C; 474 G; 305 T;

Query Match 2.4%; Score 280.4; DB 1; Length 1673;
 Best Local Similarity 99.6%; Pred. No. 7e-56;
 Matches 281; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6330 acacagctgtatcagccccagagccccaccccttcctcagctccctccgaagcta 6389
 |||||||
 DB 249 ACACAGCTGTATCAGCCCCAGAGCCCCACCCCTTCTCATCGCTCCCTGCAAGCTA 308
 |||||||
 QY 6390 cctgtctatacatgagacacacctgaggccacgcctgagggctctactgagacctta 6449
 |||||||
 DB 309 CCTGCTCTATACATGAGACACACCTGGGGCCACCGCTGAGGGGCTCTACTGAGCCCTCA 368
 |||||||
 QY 6450 atgtgtcgcgcctcctgagctgtccgcctccttaaacctccacccctggccttg 6509
 |||||||
 DB 369 ATGTGTGCGCGCTGCGCTCTGAGCTGTCCGGCTCTTAACACCTCCAGCCCTGG 428
 |||||||
 QY 6510 cccctgtctaaccttaatgtgtccagagcagctcagagagacaaatctgtgttcaagccc 6569
 |||||||
 DB 429 CCTGTGCTAACCTTATGTGGTCCAGGACAGTCAAGAAACATCTGTGTGTACGCC 488
 |||||||
 QY 6570 gagacgagcagcatctgtgctgtgctcctgctctatgttgcct 6611
 |||||||
 DB 489 GAGACGGCAGCATTTCTGTGGCTGTGCTCTGCTCTATGTGTGCT 530
 |||||||

RESULT 7
 V41688
 ID V41688 standard; CDNA; 1656 BP.
 AC V41688;
 DT 26-OCT-1998 (first entry)
 DE Nucleotide sequence of the murine U4 gene.
 KW Murine; U4 protein; haematopoietin receptor superfamily;
 KW cell proliferation; immune response; antibody; cell differentiation;
 KW autoimmune disease; cancer; allergy; ds.
 OS Mus sp.
 FT Key
 FT CDS
 Location/Qualifiers
 122..1399
 /tag= a
 /product= "U4 protein"

CC a cytokinin-like receptor. Soluble Zcytors may be administered to
CC down-regulate the effects of a growth and/or maintenance factor in
CC thyroid, heart, and skeletal muscle for example to lessen the effect
CC of cardiotrophin-1 on cardiac pathologies, so preventing heart
CC enlargement. Zcytors could be used to detect cardiotrophin-1 in the
CC blood, and to discover other possible Zcytors ligands. A probe
CC comprising Zcytors DNA or RNA can be used to determine the presence
CC and integrity of the Zcytors gene on chromosome 19. Antibodies and the
CC anti-idiotypic antibody could be used to purify Zcytors and
CC therapeutically to modify Zcytors ligand effects.
SQ Sequence 1724 BP; 350 A; 550 C; 500 G; 324 T;

Query Match 2.1%; Score 248; DB 1; Length 1724;
Best Local Similarity 92.9%; Pred. No. 2.6e-48;
Matches 260; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 6330 acacagctgtatagccccagagaccaccccttcacgtcctccctgaagcta 6389
DB 283 ACACAGCTGTATAGTCCAGCCAGACCCCACTCTTCATCGATCGTCTCCCTCATGCTA 342
QY 6390 cctgctctatagagagacacacccctgagcagctgagggctctactgagacctta 6449
DB 343 CGTCTCTATATCATGAGACACACCGGGGCGCACTGCTGAGGGCTCTACTGACCTCA 402
QY 6450 atggtgcgcgcctgcccctctgagctgtccgcctcttaaacctccaccctgcccctg 6509
DB 403 ACGGCGCCCGCGCTCCCTCAGAGCTGTCGGTCTCTCAACACCTCCACCTGCGCTTG 462
QY 6510 cccgtgtaacctaatggtgtccagcagcagcagcagcagcagcagcagcagcagcagc 6569
DB 463 CCGTGGCTAACCTTAAGGTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 522
QY 6570 gaaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 6609
DB 523 GAGATGGCAGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 562

RESULT 10
ID V70894
AC V70894 standard; cDNA: 1690 BP.
DE 17-MAR-1999 (first entry)
KW cDNA encoding human Zcytors.
KW Zcytors; cytokinin-like receptor; down-regulation; growth factor;
KW maintenance factor; thyroid; heart; skeletal muscle; cardiotrophin-1;
KW cardiac pathology; heart enlargement; Zcytors ligand; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 52..1320
FT /tag= a
FT /product= zcytors

MO9849307-A1.
PD 01-MAY-1998; 008865.
PR 13-FEB-1998; US-074721.
PR 01-MAY-1997; US-045287.
PR 01-MAY-1997; US-850030.
PR 13-FEB-1998; US-023890.
PA (ZYMO) ZYMOGENETICS INC.
PI Adams RL, Foster DC, Gilbert T, Jelmberg AC, Lehnner JM.
PI Lok S, Presnell SR, Whitmore TE.
PI WPI: 99-034662/03.
DR P-PSDB: W70860.
PT New mammalian cytokinin-like receptor Zcytors - useful for, e.g.
PT down-regulating Zcytors natural ligands or detecting cardiotrophin-1
PT in blood
PS Disclosure: Page 63-66; 55pp; English.
CC The present sequence encodes a protein designated Zcytors, which is
CC a cytokinin-like receptor. Soluble Zcytors may be administered to
CC down-regulate the effects of a growth and/or maintenance factor in
CC thyroid, heart, and skeletal muscle for example to lessen the effect
CC of cardiotrophin-1 on cardiac pathologies, so preventing heart

CC enlargement. Zcytors could be used to detect cardiotrophin-1 in the
CC blood, and to discover other possible Zcytors ligands. A probe
CC comprising Zcytors DNA or RNA can be used to determine the presence
CC and integrity of the Zcytors gene on chromosome 19. Antibodies and the
CC anti-idiotypic antibody could be used to purify Zcytors and
CC therapeutically to modify Zcytors ligand effects.
SQ Sequence 1690 BP; 319 A; 592 C; 505 G; 274 T;

Query Match 1.9%; Score 221.8; DB 1; Length 1690;
Best Local Similarity 86.8%; Pred. No. 3.4e-42;
Matches 244; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 6330 acacagctgtatagccccagagaccaccccttcacgtcctccctgaagcta 6389
DB 167 ACACAGCTGTATAGTCCAGCCAGACCCCACTCTTCATCGATCGTCTCCCTCATGCTA 226
QY 6390 cctgctctatagagagacacacccctgagcagctgagggctctactgagacctta 6449
DB 227 CGTCTCTATATCATGAGACACACCGGGGCGCACTGCTGAGGGCTCTACTGACCTCA 286
QY 6450 atggtgcgcgcctgcccctctgagctgtccgcctcttaaacctccaccctgcccctg 6509
DB 287 ATGGGCGCCCGCTCCCTCAGAGCTGTCGGTCTCTCAACACCTCCACCTGCGCTTG 346
QY 6510 cccgtgtaacctaatggtgtccagcagcagcagcagcagcagcagcagcagcagcagc 6569
DB 347 CCGTGGCTAACCTTAAGGTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 406
QY 6570 gaaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 6610
DB 407 GTACCGCAGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 447

RESULT 11
ID V70895
AC V70895 standard; cDNA: 1813 BP.
DE 17-MAR-1999 (first entry)
KW cDNA encoding an allelic variant of human Zcytors.
KW Zcytors; cytokinin-like receptor; down-regulation; growth factor;
KW maintenance factor; thyroid; heart; skeletal muscle; cardiotrophin-1;
KW cardiac pathology; heart enlargement; Zcytors ligand; allelic variant;
KW ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 88..1365
FT /tag= a
FT /product= zcytors

MO9849307-A1.
PD 01-MAY-1998; 008865.
PR 13-FEB-1998; US-074721.
PR 01-MAY-1997; US-045287.
PR 01-MAY-1997; US-850030.
PR 13-FEB-1998; US-023890.
PA (ZYMO) ZYMOGENETICS INC.
PI Adams RL, Foster DC, Gilbert T, Jelmberg AC, Lehnner JM.
PI Lok S, Presnell SR, Whitmore TE.
PI WPI: 99-034662/03.
DR P-PSDB: W70861.
PT New mammalian cytokinin-like receptor Zcytors - useful for, e.g.
PT down-regulating Zcytors natural ligands or detecting cardiotrophin-1
PT in blood
PS Disclosure: Page 68-70; 55pp; English.
CC The present sequence encodes an allelic variant of protein designated
CC Zcytors, which is a cytokinin-like receptor. Soluble Zcytors may be
CC administered to down-regulate the effects of a growth and/or maintenance
CC factor in thyroid, heart, and skeletal muscle for example to lessen the
CC effect of cardiotrophin-1 on cardiac pathologies, so preventing heart
CC enlargement. Zcytors could be used to detect cardiotrophin-1 in the
CC blood, and to discover other possible Zcytors ligands. A probe
CC comprising Zcytors DNA or RNA can be used to determine the presence

CC and integrity of the Zcytor5 gene on chromosome 19. Antibodies and the anti-idiotypic antibody could be used to purify Zcytor5 and CC therapeutically to modify Zcytor5 ligand effects.

Sequence 1813 BP; 415 A; 604 C; 519 G; 275 T;
 Query Match 1.9%; Score 221.8; DB 1; Length 1813;
 Best Local Similarity 86.8%; Pred. No. 3.5e-42;
 Matches 244; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 6330 acacagctgtaataagcccccagagaccctcttcatcgccctccctgaagta 6389
 |||||||
 DB 200 ACACAGCTGTATAGTCCCGAGATCCACGCTTCTATGCGCTCCCTCGGGCCA 259
 QY 6390 cctgctatacatgtagacacacacctgaggccagctgagggtctactcaggacctca 6449
 |||||||
 DB 260 CCTCTAGTGTACAGAGAGACCCACAGAGCCAGCGGCTCTACTGAGACCTCA 319
 QY 6450 atgtgtgcgcgcctcctctgagctgtccgcctctcttaaacctccacccctggccctg 6509
 |||||||
 DB 320 ATGGGCGCCGCTCCCTCCCTAGCTCTCCGCTACTCAAGCCCTCCACCTTGCTGG 379
 QY 6510 cccctgtaaccttaagtggtccagcagcagctcagagagaacatctgtgtacgccc 6569
 |||||||
 DB 380 CCTGGCCCAACTCAATGGGTCCAGGACGCGTGGGGGACACCTGTGTGACAGCC 439
 QY 6570 gagacggcagcatctgtgctgctcctcctctatgttggc 6610
 |||||||
 DB 440 GTGACGGCAGCATCTGTGCTGCTCTGCTCTATGTTGGC 480

RESULT 12

V41689 ID V41689 standard; cDNA; 1579 BP.

AC V41689; 26-OCT-1998 (first entry)
 DE Nucleotide sequence of the human U4 gene.
 DT Human; U4 protein; haematopoietin receptor superfamily;
 KW cell proliferation; immune response; antibody; cell differentiation;
 KM autoimmune disease; cancer; allergy; ds.
 OS Homo sapiens.

FH Key Location/Qualifiers
 FT 1..1228
 FT CDS /tag- a
 FT /product- "U4 protein"

PN MO6831811-A1.
 PD 23-JUL-1998.
 PF 15-JAN-1998; U00334.
 PR 16-JAN-1997; US-784863.
 PA (GEMV) GENETICS INST. INC.
 PI Collins M, Donaldson DD, Neben T, Whitters M;
 DR WPI; 98-414109/35.
 DR P-PSDB; W59805.
 PT New nucleic acid encoding U4 haematopoietin receptor superfamily
 chain - potentially useful, e.g. for modulating cell proliferation
 or immune response, for treating cancer and autoimmune disease
 PS Claim 1, Page 28; 38pp; English.
 CC This is the nucleotide sequence encoding the human U4 protein from
 the haematopoietin receptor superfamily, used in the method of the
 CC invention for the modulation of cell proliferation, or the immune
 CC response. Transformed mammalian cells are used to produce recombinant
 CC U4 protein. The U4 protein is used to screen for specific binding
 CC agents, raise antibodies. It is also used as reagents for assays and
 CC as tissue markers for isolation of cognate ligands and receptors, and
 CC in pharmaceutical compositions which may modulate cell proliferation,
 CC cell differentiation, and the immune system (e.g. for treating immune
 CC deficiency, inherited or the result of infection, autoimmune diseases,
 CC cancer, and allergy).
 SO Sequence 1579 BP; 304 A; 535 C; 473 G; 267 T;

Query Match 1.9%; Score 220.2; DB 1; Length 1579;
 Best Local Similarity 86.5%; Pred. No. 7.7e-42;

Matches 243; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 6330 acacagctgtaataagcccccagagaccctcttcatcgccctccctgaagta 6389
 |||||||
 DB 75 ACACAGCTGTATAGTCCCGAGATCCACGCTTCTATGCGCTCCCTCGGGCCA 134
 QY 6390 cctgctatacatgtagacacacacctgaggccagctgagggtctactcaggacctca 6449
 |||||||
 DB 135 CCTCTAGTGTACAGAGAGACCCACAGAGCCAGCGGCTCTACTGAGACCTCA 194
 QY 6450 atgtgtgcgcgcctcctctgagctgtccgcctctcttaaacctccacccctggccctg 6509
 |||||||
 DB 195 ACGGCGCGCGCTCCCTCCCTAGCTCTCCGCTACTCAAGCCCTCCACCTTGCTGG 254
 QY 6510 cccctgtaaccttaagtggtccagcagcagctcagagagaacatctgtgtacgccc 6569
 |||||||
 DB 255 CCTGGCCCAACTCAATGGGTCCAGGACGCGTGGGGGACACCTGTGTGACAGCC 314
 QY 6570 gagacggcagcatctgtgctgctcctcctctatgttggc 6610
 |||||||
 DB 315 GTGACGGCAGCATCTGTGCTGCTCTGCTCTATGTTGGC 355

RESULT 13

V27144 ID V27144 standard; cDNA; 1391 BP.

AC V27144; 23-SEP-1998 (first entry)
 DE Nucleotide sequence of clone Hfx-66 encoding human NR6.
 DT Human; Haematopoietin receptor; cell proliferation; cell differentiation;
 KW cell survival; therapeutic; neuronal proliferation; drug screening; ss;
 KM Human.
 OS Homo sapiens.

FH Key Location/Qualifiers
 FT 1..1053
 FT CDS /tag- a
 FT /product- "Human NR6"

PN MO9811225-A2.
 PD 19-MAR-1998.
 PF 11-SEP-1997; G02479.
 PR 11-SEP-1996; AU-002246.
 PA (AMRA-) AMRAD OPERATIONS PTY LTD.
 PI (DZIE) DZIELEWSKA H E.
 PI Alexander W, Fabry L, Farley A, Hilton DJ, Kikuchi Y,
 PI Kojima T, Maeda M, Nash A, Nicola NA, Rakar S, Willison T,
 PI Zhang J;
 DR WPI; 98-260970/23.
 DR P-PSDB; W55015.
 PT New isolated haematopoietin receptor - used for developing products
 for modulating proliferation, differentiation and survival of cells,
 PT e.g. neuronal cells
 PS Claim 7; Page 102-104; 182pp; English.
 CC The NR6 gene encodes a novel Haematopoietin receptor (HR). Interaction
 CC between the novel HR and a ligand facilitates proliferation. The HR and its
 CC derivatives can be used for modulating the activity of the receptors e.g.
 CC to regulate development, maintenance or regeneration in an array of
 CC different cells and tissues in vitro and in vivo. They can be present in
 CC therapeutic cells used for modulating neuronal proliferation, differentiation
 CC and survival. The products can also be used for detection and diagnosis,
 CC e.g. for cancers or predisposition to cancers, or for drug screening.
 SO Sequence 1391 BP; 281 A; 459 C; 417 G; 234 T;

Query Match 1.6%; Score 190.8; DB 1; Length 1391;
 Best Local Similarity 73.5%; Pred. No. 5.3e-35;
 Matches 291; Conservative 0; Mismatches 37; Indels 68; Gaps 1;

QY 8604 gtgacacagagaccaccacacagctgacgctgtgagggtcctgagacacag 8663
 |||||||
 DB 469 GTGACACAGGAGACCCCGCCGAGGTGACGTGAGCCCGCTGGGGGCTGTGAGGACAG 528
 QY 8664 ctgagtgctgctggtgtctaccacacagctctcaagattctcttccaagcaagtaac 8723

INFORMATION FOR SEQ ID NO: 16:

STRANDEDNESS: single


```

NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/757,390
FILING DATE: 19910910
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Mirock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7005-030
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 8698864/9741
TELEX: 66141.PENNIE
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 1355 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 13..1008
US-07-757-390-16

Query Match          0.8%; Score 99; DB 1; Length 1355;
Best Local Similarity 78.4%; Pred. No. 6,76-15;
Matches 131; Conservative 0; Mismatches 35; Indels 1; Gaps 1;

QY 5533 gttctgattcgttgcttggcttgagcttgccttgaatcttgaacaggcgcttcctgaactagaatcc 5592
      ||| |||| | |||| ||| |||| ||| |||| ||| |||| ||| |||| ||| |||| ||| ||||
Db 1355 GTGTCATCTTTTGTGTTGGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGAGACAGGAGTTCTCTGAGT 1296

QY 5593 agtcctgctgctccttgtaactcaactcctgtgaccagcgctgccttgtaactagaatcc 5652
      ||| |||| ||| |||| ||| |||| ||| |||| ||| |||| ||| |||| ||| |||| ||| ||||
Db 1295 AGCTCTGCGCTGTC-TGAACACTCACTTATAGACCACAAGCTGCGCTCGAATCCAGAATAATCC 1237

QY 5653 gccctgctgctgccccaagtccttagattaaggtgctcacctgca 5699
      ||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 1236 TCCCCTGTCTGCATCCCAAGTGCTGRTATTAAGCGCAGCTCTACCA 1190

RESULT      8
US-08-442-282-4/c
: Sequence 4, Application US/08442282
: Patent No. 5760204
GENERAL INFORMATION:
APPLICANT: Takatsu, Kiyoshi
APPLICANT: Tomimaga, Akira
APPLICANT: Takegi, Satoshi
APPLICANT: Murata, Yoshiyuki
TITLE OF INVENTION: Human And Murine Interleukin-5 Receptor
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
```


QY 4159 tgccttgcagctatctctgtagaccagctagcctcaactcaaacctactgctct 4218
DB 3215 tgccttgcagctatctctgtagaccagctagcctcaactcaaacctactgctct 3156
QY 4219 gcttctgcagctatctctgtagaccagctagcctcaactcaaacctactgctct 4264
DB 3155 gcttctgcagctatctctgtagaccagctagcctcaactcaaacctactgctct 3110

RESULT 4
US-08-555-723B-3
Sequence 3, Application US/08555723B
Patent No. 5837534

GENERAL INFORMATION:
APPLICANT: OLSON, Eric N.
APPLICANT: LI, Li
APPLICANT: MIRANO, Joseph M.
TITLE OF INVENTION: SMOOTH MUSCLE 22, PROMOTER, GENE TRANSFER
TITLE OF INVENTION: VECTORS CONTAINING THE SAME, AND METHOD OF
TITLE OF INVENTION: USE OF THE SAME TO TARGET GENE EXPRESSION IN
TITLE OF INVENTION: ARTERIAL SMOOTH MUSCLE CELLS
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS
STREET: 2100 Pennsylvania Avenue, N.W., Suite 800
CITY: Washington, D.C.
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20037

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/555,723B
FILING DATE: 14-NOVEMBER-1995
CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: KIT, Gordon
REGISTRATION NUMBER: 30,764
REFERENCE/DOCKET NUMBER: A-6663
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 293-7060
TELEFAX: (202) 293-7860
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3892 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: genomic
HYPOTHETICAL: NO

US-08-555-723B-3

Query Match 0.8%; Score 100.2; DB 4; Length 3892;
Best Local Similarity 61.9%; Pred. No. 6,1e-15;
Matches 159; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 4000 ttgctcttgtagtgcagctccagctctgacacagaggtgtgacccacagctcc 4059
DB 3326 ttgctcttgtagtgcagctccagctctgacacagaggtgtgacccacagctcc 3385
QY 4060 ctctcttccctccctccctctctcttctgctctgagactaattttcttctctt 4119
DB 3386 tcttcttcttcttct 3445
QY 4120 ttgctcttgtagtgcagctccagctctgacacagaggtgtgacccacagctcc 4179
DB 3446 tcttcttctttagacagaggttctctctgtagccctgctcttctgagactcactctt 3505

QY 4180 gaccagctagctcctcaactcaaacctactgctctgctcttccagtgctgagcacta 4239
DB 3506 gaccagctagctcctcctcaactcaaacctactgctctgctcttccagtgctgagcacta 3565
QY 4240 aagatgtgggcccacac 4256
DB 3566 aagcgtgtgcccacac 3582

RESULT 5
US-08-686-878A-38
Sequence 38, Application US/08686878A
Patent No. 5708157

GENERAL INFORMATION:

APPLICANT: Jacobs, Kenneth
APPLICANT: McCoy, John
APPLICANT: Lavallee, Edward
APPLICANT: Racie, Lisa
APPLICANT: Merberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Evans, Cheryl
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
TITLE OF INVENTION: ENCODING THEM
NUMBER OF SEQUENCES: 71
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/686,878A
FILING DATE:
CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:
NAME: Brown, Scott A.
REGISTRATION NUMBER: 32,724
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8224
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 535 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA

US-08-686-878A-38

Query Match 0.8%; Score 99.8; DE 2; Length 535;
Best Local Similarity 82.5%; Pred. No. 2,6e-15;
Matches 113; Conservative 1; Mismatches 23; Indels 0; Gaps 0;

QY 4103 tttttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 4162
DB 339 tttgttgggttttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 458
QY 4163 ctgagcactatctttagaccaggttagacctcaaacctcaaacctactgctgct 4222
DB 459 ctaragctactctttagaccaggttagacctcaaacctcaaacctactgctgct 518
QY 4223 ttcagtgctgagcacta 4239
DB 519 cccaagtctgagcacta 535

QY 4099 ttaatttttcttcttcttcttgagcatttcgacagggtctctctgtacagccctggc 4158

Db 3275 TTTATTCAGCTTTTGCTTAATTTT

Db 3275 TTTATTCAGCTTTTGCTTAATTTT

OM nucleic - nucleic search, using sw model

Title: US-09-037-657-38
 Port of origin: 11033

Sequence: 1 gcgcgcgcgtcagtgattac.....tatgtctggtgtgggga 11832

Scoring table: IDENTITY_NUC

Searched: 176461 seqs, 45838279 residues

Database : Issued_Patents_NA: *

```
1: /cgn2_6/prodata/2/ina/5A.COMB.seq.*
2: /cgn2_6/prodata/2/ina/5B.COMB.seq.*
3: /cgn2_6/prodata/2/ina/5C.COMB.seq.*
4: /cgn2_6/prodata/2/ina/5D.COMB.seq.*
5: /cgn2_6/prodata/2/ina/pcut9.COMB.seq.*
6: /cgn2_6/prodata/2/ina/backfiles1.seq.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

No.	Score	Match	Query	Length	DB	ID	Description
1	111.6	0.9	4698	1	US-07-807-0435-5	Sequence 5, Appl1	
2	111.6	0.9	4698	1	US-08-799-8499-5	Sequence 5, Appl1	
3	103.6	0.9	3293	1	US-07-923-976-1	Sequence 1, Appl1	
4	100.2	0.8	3892	4	US-08-555-7238-3	Sequence 3, Appl1	
5	99.8	0.8	535	2	US-08-866-8788-38	Sequence 38, Appl1	
6	99	0.8	1355	1	US-07-757-390-4	Sequence 4, Appl1	
7	99	0.8	1355	1	US-07-757-390-4	Sequence 4, Appl1	
8	99	0.8	1355	3	US-08-442-282-16	Sequence 16, Appl1	
9	99	0.8	1355	3	US-08-442-282-16	Sequence 16, Appl1	
10	99	0.8	1355	3	US-08-442-282-16	Sequence 16, Appl1	
11	99	0.8	1355	3	US-08-442-282-16	Sequence 16, Appl1	
12	95	0.8	7218	2	US-08-332-463-14	Sequence 14, Appl1	
13	94.2	0.8	1808	1	US-07-757-390-2	Sequence 2, Appl1	
14	94.2	0.8	1808	1	US-07-757-390-2	Sequence 2, Appl1	
15	94.2	0.8	1808	3	US-08-442-282-15	Sequence 2, Appl1	
16	94.2	0.8	1808	3	US-08-442-282-15	Sequence 2, Appl1	
17	94.2	0.8	1808	3	US-08-442-282-15	Sequence 2, Appl1	
18	94.2	0.8	1808	3	US-08-442-282-15	Sequence 2, Appl1	
19	90.6	0.8	240	2	US-08-487-7488-6	Sequence 6, Appl1	
20	89.6	0.8	2797	4	US-08-555-7238-4	Sequence 4, Appl1	
21	87.6	0.7	5687	4	US-08-380-4033-3	Sequence 3, Appl1	
22	86.6	0.7	7218	2	US-08-332-463-14	Sequence 14, Appl1	
23	85	0.7	3383	5	PCT-US95-09008-1	Sequence 1, Appl1	
24	84.8	0.7	1753	5	PCT-US95-02576-6	Sequence 6, Appl1	
25	82	0.7	26700	2	US-08-472-217-1	Sequence 1, Appl1	
26	82	0.7	26700	4	US-08-488-199-5	Sequence 1, Appl1	
27	81.2	0.7	4016	5	PCT-US95-08334A-1	Sequence 5, Appl1	
28	79.6	0.7	1179	4	US-08-465-794-4	Sequence 1, Appl1	
29	77.4	0.7	3963	4	US-08-464-961-1	Sequence 4, Appl1	
30	77.4	0.7	3963	5	PCT-US96-08233-1	Sequence 1, Appl1	
31	77	0.7	5353	4	US-08-290-733C-3	Sequence 3, Appl1	
32	74.8	0.6	3388	4	US-08-620-694A-1	Sequence 1, Appl1	
33	73.6	0.6	200	1	US-08-322-177A-8	Sequence 8, Appl1	
34	73.4	0.6	1919	1	US-07-991-587A-1	Sequence 1, Appl1	
35	73.4	0.6	1919	1	US-08-309-985-1	Sequence 1, Appl1	
36	73.4	0.6	2738	4	US-08-395-866-17	Sequence 17, Appl1	
37	72.6	0.6	4157	4	US-08-871-266B-1	Sequence 1, Appl1	

[illegible]

ALIGNMENTS

RESULT 1
US-07-807-043B-5
; Sequence 5, Application US/07807043B
; Date of Waiver 03/03/74

GENERAL INFORMATION:

APPLICANT: Boon, Thlerry, Van den Eynde, Bano t
TITLE OF INVENTION: Tumor Rejection Antigen Precursors, Tumor

NUMBER OF SEQUENCES: 16

```

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch

```

CITY: New York

ZIP: 10022

MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage

OPERATING SYSTEM

CURRENT APPLICATION DAT

FILING DATE: 19911212

PRIOR APPLICATION DATA

FILING DATE: 23-SEPTEMBER-1991

APPLICATION NUMBER: 07/728-838

FILED DATE: 3-20-11-1991

APPLICATION NUMBER: 07/105,102
FILING DATE: 23-May-1991

ATTORNEY/AGENT INFORMATION:
NAME: HANSON NO 5343774

REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: ITD 353 3

TELECOMMUNICATION INFORMATION: 633-600-8200

TELEFAX: (212) 838-3884

SEQUENCE CHARACTERISTICS:

TYPE: NUCLEIC ACID

TOPOLOGY: linear

US-07-807-043B-5

Best Local Similarity 81.68; Pred. No. 1.1e-17;

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

[illegible]

4210

```

;      TOPOLOGY: linear
;      IMMEDIATE SOURCE:
;      CLONE: PTZgpt-Fls
US-08-232-463-14

```

```

Query Match      0.8%; Score 95; DB 2; Length 7218;
Best Local Similarity 3.6%; Pred. No. 1,6e-13;
Matches 14; Conservative 254; Mismatches 119; Indels 0; Gaps 0;

```

[illegible]

RESULT 13
US-07-757-390-2/c
Sequence 2, Application US/07757390
Patent No. 5453491
GENERAL INFORMATION:
APPLICANT: Takatsu, Kiyoshi
APPLICANT: Tomioka, Akira
APPLICANT: Takagi, Satoshi
APPLICANT: Murata, Yoshiyuki
TITLE OF INVENTION: Human And Murine Interleukin-5 Receptor
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennile & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/757,390
FILING DATE: 19910910
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Mastrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7005-030
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 869864/9741

```

?       TELEX: 66141 PENNIE
?       INFORMATION FOR SEQ ID NO: 2
?       SEQUENCE CHARACTERISTICS:
?       LENGTH: 1808 base pairs
?       TYPE: NUCLEIC ACID
?       STRANDEDNESS: single
?       TOPOLOGY: linear
?       MOLECULE TYPE: CDNA
?
US-07-757-390-2

```

Query Match	0.88;	Score 94.2;	DB 1;	Length 1808;
Best Local Similarity	80.88;	Pred. No. 1.2e-13;		
Matches 122;	Conservative 0;	Mismatches 28;	Indels 1;	Gaps 1;

[illegible]

RESULT 14
 US-07-757-390-15/c
 Sequence 15 Application US/07757390
 Patent No. 5453491
 GENERAL INFORMATION:
 APPLICANT: Takatsu, Kiyoshi
 APPLICANT: Tomiatsu, Akira
 APPLICANT: Takegaki, Satoshi
 APPLICANT: Murata, Yoshiyuki
 TITLE OF INVENTION: Human And Murine Interleukin-5 Receptor
 NUMBER OF SEQUENCES: 18
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Pennie & Edmonds
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 10036-2711
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/757,390
 FILING DATE: 19910910
 CLASSIFICATION: 530
 ATTORNEY/AGENT INFORMATION:
 NAME: Mistrock, S. Leslie
 REGISTRATION NUMBER: 18,872
 REFERENCE/DOCKET NUMBER: 7005-030
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212 790-9090
 TELEFAX: 212 8698864/9741
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 15:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1808 base pairs
 TYPE: NUCLEIC ACID
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 303..1547

US-07-757-390-15

Query Match	0.88;	Score 94.2;	DB 1;	Length 1808;
Best Local Similarity	80.88;	Pred. No. 1.2e-13;		
Matches 122; Conservative	0;	Mismatches 28;	Indels 1;	Gaps 1;

DQ 5549 gttggttgagatttggtaattcagacaggcttcctctgtgaacctgctgtccatt 5608
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1808 GTTGTGTTGTTGTTGTTGTTTCAGACAGCGGTTCTCTGTGTACCTCTGGGTGTCC-T 1750

QY	5669	caagtgcctagattaaagtgctgcactgcc	56999
Dp	1689	CAGGTGCTGTATTAAAGGCGAGCTTACCA	16599

[illegible]

Query Match 0.88; Score 94.2; DB 3; Length 1808;
Best Local Similarity 80.8%; Pred. No. 1.2e-13;
Matches 122; Conservative 0; Mismatches 20; Indels 1; Gaps 1

FEATURES

Query Match	1.9%	Score 219.2	DB 45	Length 431
Best Local Similarity	86.4%	Pred. No. 1.7e-37		
Matches 242	Conservative 0	Mismatches 38	Indels 0	Gaps 0

RESULT 11
AT 70100

LOCUS 385 bp mRNA Est 14-MAY-1999
DEFINITION we65f03.x1 Soares-Thymus.NHFF.Homo sapiens CDNA clone
IMAGE:2345961 3' similar to TR:0754562 0754562 CYTOKINE-LIKE FACTOR-1-
PRECURSOR. ; mRNA sequence.
ACCESSION
NID A1670108
94834882
A1670108.1 GI:4834882
VERSION
KEYWORDS EST.

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nlh.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Seq primer: 40UP from GIBCO
High quality sequence stop: 372.

FEATURES
Source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="2"
/clone="IMAGE:2345981"
/clone_1lb="Soares-Chymus_NH7th"
/dev_stage="fetal"
/lab_host="DH10B (phage-resistant)"
/note="Organ: thymus, pooled; Vector: pUT3D-Pac
(Pharmacia) with a modified polylinker; Site_1: Not I;
Site_2: Eco RI; 1st strand cDNA was primed with a Not I -
oligo(dT) primer [5'
TGTACCACTGATGAGGGGCGCCGCAACGTTTTTTTTTTTTTTTTTTT
3'] ,
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia) , digested with Not I and cloned into the Not I
and Eco RI sites of the modified pUT3D vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M. Fatima Bonaudo. "

```

Query Match	1.8%	Score 215.4	DB 50	Length 385
Best Local Similarity	85.4%	Pred. No. 1.1e-36		
Matches 240	Conservative 0	Mismatches 41	Indels 0	Gaps 0
QY	6330	accacgctgaatcagccccacgagccccacccttcacatcgctctctccctcaagcta	6389	
DB	75	ACACACTGTGTACGTCCCAAGATCCCAAGCTTCTCATAGGCTCTCTCCCTGTGGCCA	134	
QY	6390	cctgcctctaatatgagagacacacttgggacacgccttgatgggctctacttgaactta	6449	
DB	135	CCGCGTCATGACAGGAGACCCACGAGGACCGCCGAGGGGCTTAACTGACACCTTA	194	
QY	6450	atgtgtgcgcgctgcctcttgagtgttccgctccttaaacctccaccccttggccttg	6509	
DB	195	ACGGGGCGCGCTGCCCTCGTAGCTCTCCCTGTACTCAACGCTCCACTTGTTGG	254	
QY	6510	ccctgtctaaccttaatgtgttcacgagcaagtaagagacaaatctgtgtgttcacgccc	6569	
DB	255	CCCTGGCCAACTCATATGGGTCTCAGGCGACGGTGGGGGGAACCTTGCTGCCAGGCC	314	
QY	6570	gagaagcgagcatcttgctgtgctctctccctcattgtttgc	6610	
DB	315	GTATCGGCAAGACATCCTGGCTGGCTCTCGGCTCTTATGTATGGC	355	

RESULT 12

LOCUS	482 bp	MRNA	EST	14-JUN-1996
DEFINITION	m66776.1 Scars mouse embryo NME13.5 14.5 Mus musculus cDNA			
ACCESSION	clone IMAGE:387741 5' similar to PIR.H36252 B38252 granulocyte colony-stimulating factor receptor precursor ; mRNA sequence.			
W66776				

NID	g1375694
VERSION	665776.1
KEYWORDS	GI:1375694
SOURCE	EST.
ORGANISM	house mouse. Mus musculus.
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 482)
AUTHORS	Marras,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE	The Mashu-HMI Mouse EST Project
JOURNAL	Unpublished (1996)
COMMENT	On Apr 14, 1993 this sequence version replaced g1:785250.

Contact: Maria M/Mouse EST Project
 WashU-BHMT Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 850L, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LNL ; contact the
 IMAGS Consortium (info@image.llnl.gov) for further information.
 MGI:739573
 Possible reversed clone; similarity on wrong strand
 Seq primer: EPrimer
 High quality sequence stop: 359.
 Location/Qualifiers
 1..482
FEATURES
 source

BASE COUNT	96 a	138 c	147 g	101 t
ORIGIN				

Query Match	1.8%;	Score 214.8;	DB 26;	Length 482;
Best Local Similarity	81.1%;	Pred. No. 1.6e-36;		
Matches · 305;	Conservative	0;	Mismatches 2;	Indels 69;
				Gaps 2

QY	8604	gtgcacacggagaccccccaacccagacgtagacggttgaggggcgtgagagacag	8663
Db	317	gtgaccacggagacccccccaccacccgacgtagacggttgaggggcgtgagagacag	258
QY	8664	ctgagtgtagcgtctgaggtctcacaccacagctctcaaggaattctcttcacagccaaagrac	8723
Db	257	ctgagtgtagcgtctgaggtctcacaccacagctctcaaggaattctcttcacagccaaagrac	198
QY	8724	caa-atccgcctaacccgcgttgtagagacagcgttgagcttgaaagtgccgcgtccgcgcccgac	8782
Db	197	cagatcttcgctaccgccgttgagagacagcgttgagacttgcacttgcga-----	159

[illegible]

RESULT	13
AV032198	
LOCUS	262 bp mRNA EST
DEFINITION	AV032198 Mus musculus adult C57BL/6J cerebellum Mus musculus CDNA
ACCESSION	C10E150004FF20, mRNA sequence.
NID	AV032198
VERSION	G4831746
KEYWORDS	EST.
SOURCE	house mouse.
ORGANISM	Mus musculus

REFERENCE	AUTHORS
1 (bases 1 to 262)	
Garnicini,P., Shlibata,K., Ozawa,Y., Konno,H., Ito,M., Aizawa,K., Akahira,S., Akiyama,J., Fukuda,S., Fukunishi,Y., Furugama,T., Hara,A., Hayatsu,N., Hori,F., Ishikawa,T., Itoh,M., Itawa,M., Kawai,T., Kikuchi,N., Kojima,Y., Matsuyama,T., Nitsuna,H., Oda,H., Owa,C., Sato,K., Shidata,Y., Shigemoto,Y., Shiraki,T., Sogabe,Y., Sugahara,Y., Suzuki,H., Suzuki,H., Tatemoto,M., Tomaru,Y., Tomihara,M., Watanabe,S., Yagane,M., Yamamura,T., Yokota,T., Yoshino,M., Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.	
TITLE	RIKEN MOUSE ESTS
JOURNAL	Unpublished (1999)
COMMENT	On Jun 22, 1998 this sequence version replaced gi:3246782.

Contact: Chie Owa
Genome Science Laboratory
RIKEN
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-9145
Fax: 81-298-36-9098
Email: genome-resetc.riken.go.jp
Thermosensitization and thermocitivation of thermolabile enzymes
trehalose and its application for the synthesis of full length cd
(Proc. Natl. Acad. Sci. U.S.A. 95(2):520-524 (1998))
Transcriptional sequencing: A method for DNA sequencing using RNA
polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3435-3460 (1998))
Please visit our web site (<http://genome.riken.go.jp>) for
further details.

FEATURES	Location/Qualifiers
source	1. .262
	/organism="Mus musculus"
	/strain="C57BL/6J"
	/db_xref="taxon:10090"
	/clone_1b="1500034F20"
	/clone_1b="Mus musculus adult C57BL/6J cerebellum
	/sex="male"
	/tissue_type="cerebellum"
	/dev_stage="adult"
BASE COUNT	63 a 63 c 69 g 67 t
ORIGIN	

Query Match 1.8%; Score 21.3; DB 50; Length 262;

OM nucleic - nucleic search, using sw model

Prea. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Descriptor

AA039053 n

AI421423 f

AI185780 c

Result	No.	Score	Query	Match	Length	DB	ID	Description
C	1	248	2.1	445	27	AA039053	AA039053 m199d07.t	
C	2	221.8	1.9	462	45	AI3994468	AI3994468 tf9d12.t	
C	3	221.8	1.9	474	46	AI4241243	AI4241243 tf25n01.x	
C	4	220.2	1.9	466	43	AI1610022	AI1610022 cpb9604.x	
C	5	220.2	1.9	466	43	AI1857880	AI1857880 g64d4n04.x	
C	6	220.2	1.9	477	43	AI185924	AI185924 g650c05.x	
C	7	220.2	1.9	447	43	AI185924	AI185924 g638d03.s	
C	8	220.2	1.9	417	43	AI265388	AI265388 q126b05.x	
C	9	219.2	1.9	431	48	AI579568	AI579568 ut-R-g0-u	
C	10	219.2	1.9	435	45	AI333812	AI333812 qp93612.x	
C	11	215.4	1.8	385	50	AI670108	AI670108 w65f03.x	
C	12	214.8	1.8	482	26	w65776	w65776 mel1b1.t	
C	13	213	1.8	262	50	AA033198	AA033198 AV0332198	
C	14	207.2	1.8	319	48	AI574687	AI574687 ut-R-g0-u	
C	15	200.2	1.7	390	26	W17583	W17583 mb75b01.t	
C	16	195.6	1.7	410	27	AA042914	AA042914 zk56f01.s	
C	17	176.8	1.5	273	23	R87407	R87407 ym88d09.s1	
C	18	174.8	1.5	464	29	AA049280	AA049280 m145d02.t	
C	19	171	1.4	428	45	AA866388	AA866388 ut-R-A0-e	
C	20	169.4	1.4	229	49	AA010798	AA010798 AV010798	
C	21	164	1.4	503	29	AA049278	AA049278 m145c04.t	
C	22	160.4	1.4	455	26	w46604	w46604 zc3h10.s1	
C	23	159.2	1.4	465	46	AA925924	AA925924 ut-R-A1-e	
C	24	151	1.3	470	43	AI233311	AI233311 EST229998	
C	25	143.8	1.2	167	49	AV011309	AV011309 AV011309	
C	26	135.8	1.1	464	41	AI071408	AI071408 ut-R-C2-r	
C	27	134.8	1.1	235	48	AI575060	AI575060 ut-R-g0-u	
C	28	126.8	1.1	496	44	AI318648	AI318648 u135d03.y	
C	29	126.4	1.1	370	48	AI608946	AI608946 v35605.x	
C	30	124.8	1.1	397	33	AA390114	AA390114 v28d01.t	
C	31	124	1.0	422	46	AI454475	AI454475 nm442d09.x	
C	32	122.4	1.0	365	45	AI651062	AI651062 m81609.x	
C	33	122	1.0	476	36	C79500	C79500 C79500 Mous	
C	34	122	1.0	482	47	AI509657	AI509657 vx18g08.t	
C	35	121.6	1.0	401	47	AI465880	AI465880 vx98d12.y	
C	36	121.4	1.0	458	27	AA043001	AA043001 zk56f01.t	
C	37	121.4	1.0	587	36	C79507	C79507 C79507 Mous	
C	38	121.4	1.0	309	48	AI597540	AI597540 v675b01.y	
C	39	121	1.0	422	37	AA683670	AA683670 v703e10.t	
C	40	120.8	1.0	800	50	A0067650	A0067650 A0067650	
C	41	120.4	1.0	401	39	AA881428	AA881428 vx18g08.r	
C	42	119.8	1.0	321	36	C79084	C79084 C79084 Mous	
C	43	119	1.0	398	29	AA153045	AA153045 mq53g08.r	
C	44	119	1.0	389	46	AI428461	AI428461 m303f09.x	
C	45	118.8	1.0	394	30	AA239691	AA239691 mw23h01.t	

KEYWORDS	EST.
SOURCE	house mouse.
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Crinifera; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS	1 (bases 1 to 445) Matta,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Giesel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE	The WashU-HHMI Mouse EST Project
JOURNAL	Unpublished (1996)
COMMENT	

CONTACT: Maria M/Mouse EST Project
WashU-HMI Mouse ESR Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:285477
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 441.
Location/Qualifiers
L...445

```

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone_image="IMAGE:474733"
/clone_lib="Soares mouse embryo NDM13.5 14.5"
/sex="unknown"
/tissue_type="embryo"
/dev_stage="13.5-14.5dpc total fetus"
/lab_host="DH10B"
/notes="Vector: pRT3D-Pac (Pharmacia) with a modified
polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5',
TGTTACCAATCTGATGGAGACGCGCGGAAATTTTATTTTATTTTATTTT
T 3 3], on equal amounts of mRNA from 2 13.5dpc and 2
14.5dpc embryos [total RNA provided by Minoru Ko, Wayne
State Univ., from 2 3] double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pRT3D vector. Library went through one round of
normalization, and was constructed by Bento Soares and
M.Fátima Bonaldo."

```

Query Match	2.1%	Score 248;	DB 27;	Length 445;
Best Local Similarity	99.6%	Pred. No. 9,4e-44;		
Matches 259;	Conservative	0;	Mismatches 0;	Indels 1;
			Gaps	1;
QY 11534	agctctcgcggtcaaacctcaagatagagccatctctctggttcagacctggaagc		11593	
Db 265	AGGTCCTCCGGCTAAACTTAAGATAGGCGCATCCTCTGCTGGTCAGACCTGAGGC		206	
QY 11594	tcacctgaattggagccctctgtacacctctggggaagaagaacctcaacgaagct-g		11652	
Db 205	TCACCTGAATTGGAGCCCTCTGTACACTCTGGGGCAAAAGAAACCTAACAGAGGCTGG		146	
QY 11653	ggcacaatgagctccacaacacacagcttctggtccacatgtaatggtcacactgatatac		11712	
Db 145	GGCAACAATGAGCTCCACAACACACAGCTTGTGTCCACATGATGTCACACTTGGAATAC		86	
QY 11713	cccaagtgtgggttaaggttgggtatttgaagggtctcccaagaagctctcttaataataataa		11772	
Db 85	CCCAAGTGTGGGTAGGTTGGGTAATTGCAAGGCGCTCCCAAGAAGTCTCTTTAAATAAATAA		26	

```
QY 11773 aggaagtgttcagtcacca 11792
      |||
Db 25 AGGAGTGTTCAGGTCCTCCA 6
```

RESULT	2
AI394468	
LOCUS	AI394468
DEFINITION	tF79d12.r1 NCI_CGAP_Brn23 Homo sapiens cDNA clone IMAGE:2105495 3'
ACCESSION	AI394468
NID	94224015
VERSION	AI394468.1 GI:4224015
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominae; Homo.
1 (bases 1 to 462)
NCI/NINDS-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute / National Institute of Neurological
Disorders and Stroke, Brain Tumor Genome Anatomy Project
(CGAP/RTGAP). Tumor Gene Index
Unpublished. (1998)
On Feb 17, 1998 this sequence version replaced gi:2887603.

Contact: Robert Strausberg, Ph.D.
Tel.: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: www-bio.lnl.gov/bdrg/image/image.html

```

Insert Length: 1631      Std Error: 0.00
Seq primer: -400P from GIBCO
High quality sequence stop: 454.
Location/Qualifiers
1. .462
FEATURES
source
```

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/map:10: 1"
/clone="IMAGE:2105495"
/clone_1b="NC1_CGAP_Brn23"
/tissue_type="glioblastoma (pooled)"
/lab_host="DH10B"
note="Organ: brain; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; site_1: Not I; Site-2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCGTAGGCGGCGGCGGCGGCAATCTTTTTTTTTTTTTTTTTTT
T 3']: double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT73 vector.
Library is normalized, and was constructed by Bento
Soares and M.Fatima Bonaldo."

```

Query Match	Best Local Similarity	1.9%	Score 221.8	DB 45	Length 462
Matches 244	Conservative	0	Mismatches 37	Indels 0	Gaps 0
QY 6330	acaaagctgtcaatcagccccagagccccacacctcttcacgtgtctctctccgcaagcta	6389			
Db 54	ACCAAGCTGTGATATAGTCCCAAGATCCCAAGCTTCTCATCGCTCTCCCTCTGTGGCCA	113			
QY 6390	cctgtctataacatgaggaacacacactgtgggagccacgcgtttgaggggtctactatggaacttca	6449			

	RESULT	3
LOCUS	A1421423	
DEFINITION	A1421423 474 bp mRNA EST 30-MAR-1999	
ACCESSION	U00560 Homo sapiens cDNA clone IMAGE:2097265 3	
NID	A1421423	
VERSION	94267354	
KEYWORDS	A1421423.1 GI:4267354	
SOURCE	EST.	
ORGANISM	human.	
REFERENCE	Homo sapiens	
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;	
TITLE	Eutheria; Primates; Catarrhini; Homnidae; Homo.	
	1 (bases 1 to 474)	
	NCI/NIHNS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .	
	National Cancer Institute / National Institute of Neurological	
	Disorders and Stroke, Brain Tumor Genome Anatomy Project	
	(CGAP/BRGAP), Tumor Gene Index	
	Unpublished (1998)	
COMMENT	On Apr 7, 1998 this sequence version replaced gi:3034955.	

FEATURES
SOURCE

```

length: 1664   Std Error: 0.00
primer: -40up from Gtco
quality sequence stop: 450.
Location/Qualifiers
1..474
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2097265"
/clone_1lb="NCI CGAP Brn23"
/tissue_type="gliblastoma (pooled)"
/lab_host="DH10B"
/note="Organ: brain; Vector: pTR73D-Pac (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
stranded cDNA was primed with a Not I - oligo(dT) primer [5'-
TGTTCACCAACTGTGAAGGGGCGGCGGCACATCTTTTTTTTTTTTTTTTTTTT
T 3']"; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of
the modified pTR73 vector.
Library is normalized, and was constructed by Bento
 Soares and M.Fátima Bonaldo."
181 a 181 c 128 g 83 t 1 others

```

Query Match	1.9%	Score 221.8	D3 46	Length 474
Best Local Similarity	86.8%	Pred. No. 4,7e-38		
Matches 244	Conservative 0	Mismatches 37	Indels 0	Gaps
QY 6330	acacagctgtatcatcagcccccagacccaccccttcatacgtgcctcctccgaagcta	6389		
Db 54	ACACAGCGTGTATCATAGTCCCGAGAGATCCACGCTTCTCATCGCGCTCTCCTCGTGGCCA	113		
QY 6330	ccctgcctatacatatgagacacacacctggygcacacgctgtaagggctctactggaacttca	6449		
Db 114	CGTCTCATGTGACAGGAGACCCACAGAGAGCCACCGCGAGGGCCGTACTATGACCTCCA	173		
QY 6450	atgtgtcgcgcctgcctcctcgtgagctgtcccgccctccttaaacacctcaacctgyccttg	6509		
Db 174	ATGGGCGCCCGCTGGCCCCCTCGAGCTCTCCCGTATCTCAACGCCCTCCACCTTGCGCTGG	233		
QY 6510	ccctgcgttaaccttaatggtgtccagggcagcagcaatgagbacaactgtgtgtatcagccc	6569		
Db 234	CCCTGGGCAACCTCAAGGGGTCCAGGACACGGCGGGGACAAACCTCGTGTCCACGCGCC	293		
QY 6570	gagacgacagcattctgctcgtgctcctgctcctatgatttgc	6610		
Db 294	GTGACGCGACATCTGTGGCTCTGCTCTTATGATTGGC	334		

FEATURES
SOURCE

On Jan 19, 1998, this sequence version replaced gi:2287379.

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov

This clone is available royalty-free through LNL; contact the IMABE Consortium (info@lmln1.gov) for further information.
Insert Length: 1599 Std Error: 0.00
Seq Primer: -40m3 fwd. Et from Amersham
High quality sequence stop: 454.

Location/Qualifiers

1..466

/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="839C02: 821011: 4: 4p15.33-4p16.1; 4p15.33-4p16.1"
/clone="IMAGE:1705398"
/clone_lib="Sceres-fetal_heart_NBHH19W"
/sex="unknown"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: heart; Vector: pUT73 (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer (5' TGTTACCACTGGAAGGAGCGGCGCCGACATCTTTTCTTTTCTTTT 3'), double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pUT73 vector (Pharmacia). Library went through one round of

```

[5- 'TgTTACCATCTCAAGTGGAGGCGGCCGCAATTTTTTTTTTTTTTTTTT-3']
/clone_lib="Scarses_fetal_lung_NbHL19w"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: Lung; Vector: pT7SD (Pharmacia) with a
modified polylinker. Site.1: Not I; Site.2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer
[5- 'TgTTACCATCTCAAGTGGAGGCGGCCGCAATTTTTTTTTTTTTTTTTT-3']

```

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1742408"
/clone_lib="Scares_fetal_lung_NbHL19w"
/dev_stage="19 weeks"

```

BASE COUNT
ORIGIN

81 a 183 c 129 g 83 t 1 others

"/notab="DH10B (ampicillin resistant)"
/host="Organ: lung; Vector: pT73D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - Oligo(dT) primer
[5'-TGTATCCCATCTGAAGTGGAGCGGCCCATATTTTATTTT-3']
double-stranded cDNA was size selected, ligated to Eco RI
adapters (pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M.Fátima Bonaldo. This library was constructed
from the same fetus as the fetal heart library, Soares
fetal heart NDH19W."

Query Match	1.9%	Score 220.2	DB 43	Length 477
Best Local Similarity	86.5%	Pred. No. 1.1e-37		
Matches 243; Conservative		0; Mismatches 38;	Indels 0;	Gaps 0

Dd 57 ACACAGCGTGATCAGTCCGCCAGGATCCACGGCTTCATCGGCTCCGCCGTGCGCA 116

Dy 6390 cctgctcctatcacatggagacacacctgggcaccgcgtgaggcgctctaacttgtaaccttca 6443
||||||| || | ||||| |||| | ||||| ||||| ||||| ||||| |||||
Db 117 cctgctagtgcacaggagaccacccagagaccgccgagggcccttacttggaaccttca 176

Dy 6450 atgtctgcgcgcgccccttgagcctgtccgcgcctcttaacatcacaccctgcaccttg 6509
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 177 acggcgccgccctcccccttgagctctccgtactcaacgccctcacaccttggctctgg 236

qy 6510 cccttgctaacttaatggtccagggcagcagtcagagacaatctgtgtgtcacgccc 6569
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 237 ccctggccaactcaatgggtccagggcagcagtcggtgggacaaactcgtgtgccacgccc 296

DQ 6570 gagacggcagcatctctgctgcgtccctcgccctatgtttgac 6610
 | ||||| | ||||| | ||||| | ||||| | ||||| | |||||
Db 297 GTGACGGCAGCATCCGTGGCTGGCTGCCTATGTTGGC 337

RESULT 7
A1187074

DEFINITION	qes38a03.s1 Soares_fetal_lung_NbHL19W Homo sapiens cDNA clone IMAGE:1741228 3', mRNA sequence.
ACCESSION	AI187074

```
VERSION      AI187074.1  GI:3737712
KEYWORDS     EST.
SOURCE       human.
```

REFERENCE
1 (bases 1 to 447)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.

TITLE
National Cancer Institute, Cancer Genome Anatomy Project (CGAP)
Tumor Gene Index
Unpublished (1997)

ON AUG 21, 1990 THIS SEQUENCE VERSION REPLACED.

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550

Email: robert_cslausberg@linl.gov
This clone is available royalty-free through LINL ; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Insert Length: 1667 Std Error: 0.00

ES	Location/Qualifiers	High quality sequence stop: 431.	seq primer: 40m13 1w0. EJ from Amersham
urice	1-447		

/organism="homo sapiens"

```

/db_xref="taxon:9606"
/clone_image:1741228"
/clone_lib="Soares_fetal_lung_NBHL19W"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: lung; Vector: pT73D (Pharmacia) with a
modified polylinker. Site.1: Not I; Site.2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer
[5'-TGTTCCAAATCTGAATGGGAGCGCCGCAATTTTCTTTTCTTTT-3']
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). Library went through one round of
normalization to a cot = 5. Library constructed by Bento
Soares and M.Fátima Ronaldo. This library was constructed
from the same fetus as the fetal heart library, Soares
fetal heart NBHL19W."

```

Query Match	1.9%	Score 220.2	DB 43	Length 447
Best Local Similarity	86.5%	Pred. No. 1e-37		
Matches 243; Conservative	0	Mismatches 38;	Indels 0;	Gaps 0

6350 aacacagctgatacagcccccagagaccccacccctctcaatcggtctcccccctgcaagca 6389
 6390 aacacagctgatacagcccccagagaccccacccctctcaatcggtctcccccctgcaagca 6429
 6430 aacacagctgatacagcccccagagaccccacccctctcaatcggtctcccccctgcaagca 6469
 6470 aacacagctgatacagcccccagagaccccacccctctcaatcggtctcccccctgcaagca 6509
 6510 aacacagctgatacagcccccagagaccccacccctctcaatcggtctcccccctgcaagca 6549
 6550 aacacagctgatacagcccccagagaccccacccctctcaatcggtctcccccctgcaagca 6589
 6590 aacacagctgatacagcccccagagaccccacccctctcaatcggtctcccccctgcaagca 6629
 6630 aacacagctgatacagcccccagagaccccacccctctcaatcggtctcccccctgcaagca 6669
 6670 aacacagctgatacagcccccagagaccccacccctctcaatcggtctcccccctgcaagca 6709
 6710 aacacagctgatacagcccccagagaccccacccctctcaatcggtctcccccctgcaagca 6749
 6750 aacacagctgatacagcccccagagaccccacccctctcaatcggtctcccccctgcaagca 6789
 6820 aacacagctgatacagcccccagagaccccacccctctcaatcggtctcccccctgcaagca 6859
 6860 aacacagctgatacagcccccagagaccccacccctctcaatcggtctcccccctgcaagca 6900

6390 cctgctcctataacatgtgagacacacccctgggcccacgcgcgagggctcctactgtgaccttca 6449
 135 cctgctcctataacatgtgagacacacccctgggcccacgcgcgagggctcctactgtgaccttca 194

[illegible]

Qy 6510 ccctggttaaccttaatgggtcttcacgacgacgtcagtagagacaatctgtgtgtccacgcc 6569

Db 255 ccctggcccaacctcaatgggtcttcacgacgacgtcgggggacacactctgtgtccacgcc 314

Qy	6570	gagacgacagcattctgctgctccctgcctctatgtttgc	6610
Db	315	GTGACGGCAGCATCTGCTGGCTCCCTGCTTATGTGGC	355

RESULT 8
A1269388

DEFINITION	q126b05.x1 Soares_nhmPu_S1 Homo sapiens cDNA clone IMAGE:1857585
ACCESSION	U126938
	3', mRNA sequence.

```

VERSION      AI259388.1  GT:388555
KEYWORDS
SOURCE       human.

```

Eukaryote; Metazoa; Chordata; Cranata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 1 (bases 1 to 417)

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
UNPUBLISHED (1997)
JOURNAL

On Aug. 21, 1996 this sequence version replaced.

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550

Email: robert@stabsbergen.nl
 This clone is available royalty-free through ILMT ; contact the
 IMAGE Consortium (info@ilmt.llnl.gov) for further information.
 Insert length: 1634 seq. Error: 0.00

seq primer: -400P from GIDCO

oligo-dT track. The library was constructed as described by Bonaldo, Lennon and Soares, Genome Research 6:

by: Bonaldo, Lennon and Soares, Genome Research 6: 781-806 1996

BASE COUNT	46 a	116 c	88 g	68 t	1 others
ORIGIN					

Query Match	1.8%	Score 207.2;	DB 48;	Length 319;
Best Local Similarity	92.0%;	Pred. No. 6.2e-35;		
Matches 218: Conservative	0:	Mismatches 19:	Indels 0:	Gaps 0:

QY 6330 acacagctgtatcagccccagagaccaccccttcacatcagctctccctgcgaagcta 6389
 83 ACACAGCTGTATCAGTCCCGAGACCCCTCTTCATCGGATCCTCCCTCATGCTA 142

143 CGTCTCTATACATGGAGACACACCGGGGGCAGCTGTGAGGGCTCTACTGACCTCA 202
 6530 cctcctctatcatctgagacacaccttgggacacgcgtgagggctctactgacctca 6443
 QY

D6 203 ACGGCGCCGGCTGCCCTCAGAGGTGTCTNCGTCTCCAACACTCCACCCTGGGCCTTG 262
atggtagcgcacgtgcccctctgagctgtcccgccctctaacaccctcacacctgccccgtg 6509

QY 6510 cccctggcctaacccttaatggtccagcagcagtcagagagacaatctggtgtgtcaccg 6566
 |||||
 Db 263 ccctggcctaacccttaatggtccagcagcagtcagagagacaatctggtgtgtcaccg 319

RESULT 15
W17583

DEFINITION	mus musculus p33NH19.5 mus musculus cdna clone
IMAGE	IMAGE:335209 5' / mRNA sequence.
ACCESSION	M17583
DATE	19910307

ORIGIN	USA
KEYWORDS	EST.
SOURCE	house mouse.

[illegible]

Schellenberg, K., Steptoe, M., Tan, F., Underwood, R., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, P.

JOURNAL
COMMENT
Unpublished (1996)
On May 9, 1995 this sequence version replaced g1:804170

WashU-HHMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501 St Louis MO 63108

Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through RINT: contact the

MGI:216609
Seq primer: mob.REGA+ET
High quality sequence stop: 384.

```
source
1..390
/organism="Mus musculus"
/db_xref="taxon:10090"
```

```

/clone="IMAGE:335209"
/clone_1lb="Soares mouse p3NMF19.5"
/dev_stage="19.5 dpc total fetus"

```

```
/note--"vector":PI/T3D (Pharmacica) with a modified
```


polylinker; site 1: Not I; site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer (5' TGTTCACATCTGAGTGGAGCGCCGCAATTTTCTTTT 3'), double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M. Fatima Bonaldo. RNA was kindly provided by Dr. Minoru Ko (Wayne State University).

BASE COUNT
ORIGIN

89 a 111 c 114 g 76 t

Query Match 1.7%; Score 200.2; DB 26; Length 390;
Best Local Similarity 97.4%; Pred. No. 2.2e-33;
Matches 225; Conservative 0; Mismatches 3; Indels 3; Gaps 2;

QY 11534 aggtcctgcccgttaacttaagatagagcatcctcctgctggttcagacctgagagc 11593
|||||
DB 160 AGGTCTGCGCGCTAAACTTAAGATAGGCCATCTCTGCTGSGGTCAAGACCTGAGAGC 219
QY 11594 tcacctgaattgagagccctctgtaccatctgggcaacaagaacctaccagaggt-g 11652
|||||
DB 220 TCACCTGAATTGGAGCCCTCTGTACCATCTGGGCAACAAGAACTTACAGAGGCTGG 279
QY 11653 ggcacaatagctcccaacaacacagcttgggtccacatgatggttcaacacttgatatatc 11712
|||||
DB 280 GGCACATGAGCTCCACACACACAGCTTTGGTCCACATGATGTCACACTTGATATATC 339
QY 11713 cccagtg--tggttagaggttggtatctgcaaggcctcccaaggtctctt 11761
|||||
DB 340 CCCAGTGTGGTGAAGGTTGGGCTATTGCAAGGCTCCCAAGAGTCTCTT 390

Search completed: September 19, 1999, 15:16:34
Job time: 13168 sec

PR 03-FEB-1989; US-306503.
 PA (GENE-) GENETICS INST INC.
 PA (WIT-) WHITEHEAD INST.
 PI D'Andrea A, Mong G;
 DR WPI: 90-260931/34.
 DR N-PSDB; 005748.
 PT Erythropoietin receptor and gene - used for developing reagents
 PS and systems to control and study erythropoiesis.
 PT Disclosure; Fig 2; 53pp; English.
 CC The sequence was deduced from DNA obd. from a clone isolated from
 CC a commercially available human genomic cDNA library in phage
 CC Lambda fix (Stratagene). The sequence encodes a type I trans-
 CC membrane protein with binding affinity for EPO. The gene and
 CC recombinant EPO receptor produced on expression of the DNA are
 CC used to develop reagents and systems to control and study
 CC erythropoiesis. It is believed that the EPO receptor is dys-
 CC functional in individuals with Diamond Blackfan anaemia, and may
 CC be hyperactive in polycythemia vera.
 CC See also R06511 (murine EPO receptor).
 SQ Sequence 508 AA;

Query Match 96.8%; Score 30; DB 1; Length 508;
 Best Local Similarity 80.0%; Pred. No. 4.4e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MSXMS 5
 DB 233 WSAMS 237

RESULT 3
 ID R24017
 AC R24017 standard; Protein; 552 AA.
 DT 26-NOV-1992 (first entry)
 DE Fusion protein GM-CSFRFC.
 DE Granulocyte macrophage-colony stimulating factor; GM-CSF; IgG1;
 KW Immunoglobulin G1.
 OS Synthetic.
 FH Key
 FT region Location/Qualifiers
 FT 1..318
 FT /note="human GM-CSF receptor"
 FT 319..336
 FT /note="linker and hinge"
 FT 337..445
 FT /note="IgG1 CH2"
 FT 446..552
 FT /note="IgG1 CH3"
 FT region
 PN EP-488170-A.
 PD 03-JUN-1992.
 PF 26-NOV-1991; 120187.
 PR 28-NOV-1990; DE-037837.
 PA (BEHN) BEHRINGER AG.
 PI Lauffer L, Oquendo P, Zettlmeisel G;
 DR WPI: 92-185084/23.
 PT Cell free receptor binding test contg. recombinant fusion protein -
 PT comprising carrier bound to fusion partner coupled to fixed support,
 PT and second, labelled binding partner, for receptor or antibody
 PT screening etc.
 CC Example; Fig 12; 24pp; German.
 CC The sequence is that of fusion protein GM-CSFRFC comprising the
 CC extracellular domain of granulocyte macrophage-colony stimulating
 CC factor (GM-CSF) fused via a hinge region to the FC part of the heavy
 CC chain of human IgG1. It may be used as part of a cell free receptor
 CC binding test which can be used for the identification of agonists,
 CC antagonists, antibodies, biological activity of soluble cellular
 CC receptors, functional analysis of modified ligands and diagnostic or
 CC therapeutic substances. See also R24016.
 SQ Sequence 552 AA;

Query Match 96.8%; Score 30; DB 1; Length 552;

Best Local Similarity 80.0%; Pred. No. 4.8e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MSXMS 5
 DB 306 WSSMS 310

RESULT 4
 ID R10795
 AC R10795 standard; Protein; 622 AA.
 DT 25-APR-1991 (first entry)
 DE Human prolactin receptor.
 KW Human prolactin receptor; PRL; assay; antibody; growth hormone.
 OS Homo sapiens.
 FH Key
 FT peptide Location/Qualifiers
 FT 1..24
 FT /label="sig-peptide"
 FT 25..622
 FT /label="mat-protein"
 FT protein
 PN US492378-A.
 PD 12-FEB-1991.
 PF 16-DEC-1988; 286445.
 PR 16-DEC-1988; US-286445.
 PA (ROYA-) ROYAL INST ADVAN LE.
 PI Kelly PA, Dillane J;
 DR WPI: 91-063441/09.
 DR N-PSDB; Q10550.
 PT Isolated cDNA sequence encoding human prolactin receptor - useful
 PT for expressing the receptor, e.g. for screening assays and antibody
 PT prodn.
 PS Disclosure; Fig. 1 (A-E); 11pp; English.
 CC The human PRL receptor cDNA is isolated by screening a lambda gt 10
 CC library prepared from normal human hepatoma Hep G2 and T47-D breast
 CC cancer cells. Initially, 1x10⁶ recombinants are screened with the
 CC following probe: a complementary RNA of the P3 cDNA of the rat
 CC Hep G2 library, which was later used as a probe to rescreen the
 CC library. Five additional cDNAs were identified, and by combining
 CC two of the six cDNAs, a PRL receptor cDNA of 2.5 kb was constructed,
 CC contg. a single ORF of 1866 bp. Similar partial length cDNA were
 CC isolated from the T47-D library.
 CC Several regions of sequence identity between the human growth
 CC hormone and PRL receptors can be found, both in the extracellular
 CC and cytoplasmic domains.
 CC The sequence is an important genetic engineering tool which may be
 CC used for the screening of growth hormone variants, for the development
 CC of test kits to measure PRL receptor levels in human breast and prostate
 CC cancer biopsies, for the measurement of bioactive forms of prolactin,
 CC and for the development of drugs to induce stimulation or inhibition of
 CC the immune system.
 SQ Sequence 622 AA;

Query Match 96.8%; Score 30; DB 1; Length 622;
 Best Local Similarity 80.0%; Pred. No. 5.4e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MSXMS 5
 DB 215 WSAMS 219

RESULT 5
 ID R10919
 AC R10919 standard; Protein; 400 AA.
 DT 08-MAY-1991 (first entry)
 DE Human GM-CSF receptor.
 KW Granulocyte-macrophage colony-stimulating factor; myeloid leukaemia;
 KW autoimmune disease.
 FH Key
 FT Location/Qualifiers

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: September 17, 1999, 03:03:09 : Search time 64.1 Seconds
(without alignments)

1.848 Million cell updates/sec

Title: US-09-037-657-1

Perfect score: 31

Sequence: 1 WSXWS 5

Scoring table: BLOSUM62

Searched: 188963 seqs, 23686106 residues

Database: A_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	96.8	507	1 R06511	EPO receptor seque
2	30	96.8	508	1 R06512	Epo receptor: Eryt
3	30	96.8	552	1 R24017	Fusion protein GM-
4	30	96.8	622	1 R10795	Human prolactin re
5	30	96.8	400	1 R10919	Human GM-CSF recep
6	30	96.8	211	1 R22228	Truncated human pr
7	30	96.8	211	1 R42273	Truncated human pr
8	30	96.8	284	1 R23971	MPLV env-vmpl fusi
9	30	96.8	635	1 R23970	MPLV env protein w
10	30	96.8	382	1 R27659	MPLV-env related p
11	30	96.8	184	1 R27660	Partial MPLV-env r
12	30	96.8	507	1 R47517	MEL EPO receptor.
13	30	96.8	508	1 R47518	Human EPO receptor
14	30	96.8	265	1 R50326	Mouse soluble EPO
15	30	96.8	507	1 R50327	Mouse erythropoiet
16	30	96.8	507	1 R65502	Human erythropoiet
17	30	96.8	508	1 R65503	Human erythropoiet
18	30	96.8	5	1 R68782	Thrombopoietin-der
19	30	96.8	508	1 R70032	Human erythropoiet
20	30	96.8	626	1 R75939	Human myelopoietin
21	30	96.8	635	1 R75940	Human myelopoietin
22	30	96.8	482	1 R75941	Human myelopoietin
23	30	96.8	633	1 R79908	Soluble murine MPL
24	30	96.8	633	1 R79908	Type I MPL recepto
25	30	96.8	633	1 R79908	Mouse type I MPL r
26	30	96.8	967	1 R83441	A. thaliana RPS2 po
27	30	96.8	423	1 R92813	Human interleukin
28	30	96.8	432	1 R92813	Murine interleukin
29	30	96.8	934	1 R92813	Cycloisomaltooligo
30	30	96.8	934	1 R92813	Cycloisomaltooligo
31	30	96.8	422	1 R98939	Synthetic human er
32	30	96.8	441	1 R98939	Human interleukin
33	30	96.8	441	1 R98939	Murine Etl-2 gene
34	30	96.8	633	1 R98948	Mouse type I MPL r
35	30	96.8	633	1 R98948	Mouse type I MPL r
36	30	96.8	528	1 W69363	Aeromonas caviae x
37	30	96.8	898	1 W78898	Rat UNC-5 homologu
38	30	96.8	1522	1 W93302	Human BA13 protein
39	30	96.8	1584	1 W93300	Human BA13 protein
40	29	93.5	323	1 P90527	B cell stimulating
41	29	93.5	344	1 P90528	B cell stimulating
42	29	93.5	386	1 P90526	B cell stimulating
43	29	93.5	468	1 P90525	B cell stimulating

ALIGNMENTS

44 29 93.5 810 1 R04574
45 29 93.5 218 1 W40287

Derived amino acid
Human TSP1 protein

RESULT 1
ID R06511
AC R06511
DT 04-JAN-1991 (first entry)
DE EPO receptor sequence deduced from DNA of clone 190.
KW Erythropoietin; Diamond Blackfan anaemia; polycythemia vera.
OS Mus musculus.
FT Key Location/Qualifiers
FT peptide 1..24
FT domain /label-signal peptide
FT domain 25..248
FT domain /label-extracellular domain
FT domain /note-EPO binding region
FT domain 248..271
FT domain /label-transmembrane domain
FT domain 272..507
FT domain /label-intracellular domain
FT modified_site 75..77
FT modified_site /label-N-linked_glycos
FT modified_site 182..184
FT modified_site /label-N-linked_glycos
PN W09008822-A.
PD 09-AUG-1990; U00635.
PF 01-FEB-1990; U00635.
PR 03-FEB-1989; US-306503.
PA (GENE-) GENETICS INST INC.
PA (WHIT-) WHITEHEAD INST.
PI D'Andrea A, Wong G;
DR WPT: 90-260931/34.
DR N-PSDB: Q05747.
DR Erythropoietin receptor and gene - used for developing reagents
PT and systems to control and study erythropoiesis.
PS Disclosure: fig 1; 53pp; English.
CC The sequence was deduced from DNA from a clone isolated from a
CC cDNA library prep. from uninduced murine erythroleukemia cells.
CC It is a type I transmembrane protein with binding affinity for EPO.
CC The gene and recombinant EPO receptor produced on expression of
CC the DNA are used to develop reagents and systems to control and
CC study erythropoiesis. It is believed that the EPO receptor is
CC dysfunctional in individuals with Diamond Blackfan anaemia, and
CC may be hyperactive in polycythemia vera.
CC See also R06512 (human EPO receptor).
SQ Sequence 507 AA;

Query Match 96.8%; Score 30; DB 1; Length 507;
Best local similarity 80.0%; Pred. No. 4.4e-02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WSXWS 5
DB 232 WSXWS 236
↓ Sequence Comparison "A"

RESULT 2
ID R06512
AC R06512
DT 04-JAN-1991 (first entry)
DE EPO receptor.
KW Erythropoietin; Diamond Blackfan anaemia; polycythemia vera.
OS Homo sapiens.
PN W09008822-A.
PD 09-AUG-1990.
PF 01-FEB-1990; U00635.